Methods for Remote Determination of CO_2 Emissions

Contact: Daniel McMorrow - dmcmorrow@mitre.org

January 2011

JSR-10-300

JASON
The MITRE Corporation
7515 Colshire Drive
McLean, Virginia 22102-7508
(703) 983-6997

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
January 13, 2011	Technical	

4. TITLE AND SUBTITLE

Methods for Remote Determination of CO2 Emission

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
13109022

5e. TASK NUMBER
PS

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
JASON Program Office
7515 Colshire Drive
McLean, Virginia 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

JSR-10-300

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Research & Development for National Security
Science and Technology (NA-121)
National Nuclear Security Administration
1000 Independence Avenue, SW
Washington DC 20585

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Limiter Statement A. Approved for public release. Distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

JASON was asked to assess U.S. capabilities for estimating greenhouse gas (GHG) emissions in support of monitoring international agreements. In particular, JASON was asked:

- What are the current and future capabilities for estimating GHG emissions using direct measurements of GHG atmospheric concentrations?
- How do direct measurements compare to indirect ("proxy") methods for monitoring emissions?
- What are the capabilities for estimating emissions from non-cooperating countries i.e. those for which access to in country measurements is impeded or denied?

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dimitri Kusnezov
Uncl	Uncl	Uncl	UL		19b. TELEPHONE NUMBER *(include area code)* 202-586-180

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Contents

1 EXECUTIVE SUMMARY **1**
- 1.1 Overview of the Study . . . 1
- 1.2 Direct Measurements . . . 2
- 1.3 Monitoring Energy Infrastructure (Proxy Measurements) . . . 4
- 1.4 Roadmap . . . 6

2 INTRODUCTION **9**
- 2.1 Scope of the JASON Study . . . 10
- 2.2 Greenhouse Gases, CO_2, and the Carbon Cycle . . . 12
- 2.3 Estimating CO_2 Emissions from Fossil Fuel Burning . . . 15
  - 2.3.1 Country and regional fossil-fuel CO_2 emissions . . . 17
  - 2.3.2 Benchmarks for accuracy of CO_2 emission estimates . . . 19
- 2.4 Monitoring of Energy Infrastructure as a Near-Term Alternative to Direct Measurements . . . 22

3 ENERGY INFRASTRUCTURE MONITORING **25**
- 3.1 Monitoring Supply Sectors . . . 27
  - 3.1.1 Electricity generation sector . . . 28
  - 3.1.2 Electricity distribution system . . . 31
  - 3.1.3 Fossil fuel production, distribution, and sequestration . . . 31
- 3.2 Monitoring Demand Sectors . . . 36
- 3.3 Estimating CO_2 Emissions using Knowledge of Energy Infrastructure . . . 37

4 DIRECT MEASUREMENTS **39**
- 4.1 Context . . . 40
  - 4.1.1 Measurements . . . 41
  - 4.1.2 Coordination of measurements and models . . . 44
- 4.2 In-Situ Measurements . . . 48
  - 4.2.1 Direct CO_2 measurement techniques . . . 48
  - 4.2.2 IR absorption in gas cells . . . 50
  - 4.2.3 Cavity ring-down spectroscopy (CRDS) . . . 54
  - 4.2.4 Non-dispersive infrared . . . 55
  - 4.2.5 ABET challenge . . . 58
  - 4.2.6 Sensor network . . . 59

4.2.7 Aerial . . . 62
4.2.8 Ocean . . . 63
4.3 Isotopes . . . 64
4.4 Satellite Measurements of CO_2 . . . 65
4.4.1 Introduction . . . 65
4.4.2 Measuring atmospheric CO_2 by reflected sunlight . . . 68
4.4.3 Potential for monitoring power plant emission . . . 70
4.4.4 Thermal IR measurements of atmospheric CO_2 . . . 72
4.4.5 The Merits of geosynchronous orbit . . . 75
4.4.6 L1 monitoring of CO_2 . . . 80
4.5 Findings and Recommendations: Direct Measurements . . . 81

5 MODELING/ASSIMILATION/INVERSION 85
5.1 The Flux Inversion Problem . . . 85
5.1.1 Introduction . . . 85
5.2 Discretization Issues . . . 88
5.3 Satellite Measurements . . . 93
5.4 Systematic Model Errors . . . 94
5.5 Reprise . . . 99

6 SPECIFIC FINDINGS AND RECOMMENDATIONS 101
6.1 Role of Direct Measurements . . . 101
6.1.1 Remote sensing . . . 102
6.1.2 Sensor development and deployment . . . 103
6.1.3 Isotopic measurements . . . 104
6.2 Role of Proxy Observations: Monitoring Energy System Infrastructure . . . 106
6.3 Role of Modeling . . . 108
6.4 Roadmap . . . 109

A APPENDIX: JASON and Inexpensive NDIR 115

B APPENDIX: Aircraft Monitoring of Carbon 127
B.1 Airline *in-situ* Measurements . . . 128
B.2 Open Skies Aircraft Program . . . 133
B.3 Future Directions . . . 134

C APPENDIX: Ocean Issues **137**
C.1 Air-sea CO_2 Fluxes . . . 137
C.2 Measuring CO_2 Concentrations Autonomously . . . 140
C.3 Monitoring Land Emissions from Ships . . . 142
C.4 Process Teams . . . 143

D APPENDIX: Isotopic Signatures **147**
D.1 Dilution by fossil carbon . . . 147
D.2 Carbon 14 . . . 147
D.3 Relevance to Treaty Verification . . . 151
D.4 The Suess Effect . . . 153
D.5 Measuring ^{14}C . . . 154
D.6 Stable Carbon Isotopes . . . 155
D.7 Oxygen Isotopes . . . 157
D.7.1 Oxygen Isotope Diagnostics . . . 158
D.8 Isotope Clumping . . . 159

E APPENDIX: Statistical Data Assimilation **165**
E.1 JASON's 'Carbon' Problem . . . 165
E.2 Introduction . . . 166
E.3 Saddle Path Approximation: 4DVar . . . 173
E.4 How Many Measurements are Required? . . . 174
E.5 Other Approaches to Data Assimilation . . . 177
E.6 Relation to the Carbon Problem . . . 180
E.7 Summary and Status . . . 186

F APPENDIX: Use of Tracers for Empirically Testing Models **191**
F.1 Release of Atmospheric Tracer Gases . . . 192
F.2 Continental-Scale Release Experiments in the Past . . . 193
F.3 Recommended capabilities for tracer gas release . . . 194

G APPENDIX: Selected Acronyms **197**

1 EXECUTIVE SUMMARY

1.1 Overview of the Study

JASON was asked to assess U.S. capabilities for estimating greenhouse gas (GHG) emissions in support of monitoring international agreements. In particular, JASON was asked:

- What are the current and future capabilities for estimating GHG emissions using *direct measurements* of GHG atmospheric concentrations?
- How do direct measurements compare to *indirect ("proxy") methods* for monitoring emissions?
- What are the capabilities for estimating emissions from non-cooperating countries i.e. those for which access to in-country measurements is impeded or denied?

The measurement of GHGs is a large and complex technical area. Our study focused primarily on measurement of the most prominent anthropogenic GHG component, atmospheric CO_2 arising from fossil fuel emissions. The study included the following as primary topics of investigation:

- *Direct measurements* of atmospheric CO_2, including remote observations via satellite and *in situ* measurements using land, air, or ocean-going samplers. The sub-topics included:
  - Modeling, data assimilation, and inversion needed to derive CO_2 fluxes from measured CO_2 concentrations, and to optimize measurement strategies.
  - Isotopic measurements using both stable and unstable isotopes.
- *Energy infrastructure (proxy) data* as an alternative or supplement to direct measurements.

The primary conclusions of our study are: (1) energy infrastructure monitoring provides a near-term capability for assessing a country's CO_2 emissions from fossil-fuel burning, and (2) direct measurements (including the required modeling of GHG transport in the atmosphere) will take additional development before they are capable of estimating emissions of GHGs at accuracies that are useful for supporting the monitoring of international agreements.

1.2 Direct Measurements

Background

The relevant quantities that can be measured directly are CO_2 concentrations (mole fraction in parts per million) by *in situ* observations, or CO_2 column concentrations (weighted mole fraction) looking down at a patch of the Earth from an orbiting satellite. For the purposes of treaty monitoring, what are needed are CO_2 *fluxes* (molecules per second) emitted into the atmosphere due to human activities in an area of interest. Thus the performance of a measurement system for CO_2 emissions (fluxes) depends both on the concentration measurements themselves and on what is called the "model inversion" to go from the measured concentrations to the derived emissions.

The naturally occurring carbon cycle includes CO_2 exchange to and from the atmosphere by processes in the biosphere, soils, and oceans, as well as geological processes (e.g. volcanos). Assessment of emissions due to human activities is difficult because the CO_2 emitted by anthropogenic sources is only a small fraction of the natural emission and absorption processes. A further challenge is that the transport of the CO_2 from its source to the measurement point is highly dependent on detailed meteorology, i.e. wind speed, temperature, pressure, and other variables.

Goals of Direct Measurements

As is described in the study report, JASON has chosen $\pm 20\%$ (90% confidence) measurement uncertainty as the initial benchmark by which to eval-

uate the ability of the U.S. to estimate annual anthropogenic CO_2 emission from a country or region. This benchmark was chosen as a reasonable compromise between what is we believe is achievable and what is required for treaty verification monitoring. The largest GHG emitting nations include: China, the U.S., Russia, Japan, and India, as well as the European Union as a region. For some of these, direct measurements at the $\pm 20\%$ level will be difficult to achieve as it is not only the total emission level that is important, but also the emission density (i.e. tonnes of CO_2 per square kilometer per year). We suggest starting with the larger emitters: the European Union, China, and the U.S., with the goal of developing a direct measurement capability for these at the $\pm 20\%$ level or better.

Assessment of Capabilities: Direct Measurements

- For *cooperative* countries, the technology currently exists to directly monitor GHG emissions sufficiently well on an annual basis to support U.S. decision-making on international agreements. We believe that a properly designed *in-situ* sensor network using currently available technology could reach a $\pm 20\%$ goal if augmented with observations from existing satellite instruments (e.g., AIRS, GOSAT, and SCIAMACHY) as well as accurate meteorological data. The capability will be further improved by a successful launch of the Orbiting Carbon Observatory (OCO-2) in 2013.

- For *non-cooperative* countries, there is currently no demonstrated capability to estimate country-level emissions using direct measurements of atmospheric CO_2 that has sufficient accuracy to support monitoring of compliance with international agreements.

- It is difficult to predict when direct measurements of CO_2 will yield useful emission estimates for non-cooperating countries. The situation should improve over the next 3-5 years with data from GOSAT, the possibility of surface CO_2 data from AIRS, and future data from OCO-2. Coupled with sensor networks optimized to sample downwind of

specific countries, the satellite data could provide a capability within 5 years for estimation of annually averaged net fluxes with an accuracy possibly as good as ±20%.

- Comparing remote with *in situ* measurements and national inventories in *cooperative* countries will be very instructive in deciding whether it is feasible to monitor noncooperative countries by remote measurements alone.

- Improved modeling/assimilation/inversion tools will be required to reach a goal of ±20% for both cooperative and non-cooperative countries.

- On a longer time scale (5-10 years) a GHG observation satellite in geosynchronous (or other optimized) orbit could provide a significantly improved capability for estimating GHG emissions from non-cooperating countries.

- Availability of low-cost CO_2 sensors with 1 ppm accuracy will enable *in situ* monitoring from large numbers of land, air, and sea platforms.

- Isotopic measurements will be useful as tracers of CO_2 from fossil fuel burning and other processes, though considerable work will be required to establish quantitative assessments.

1.3 Monitoring Energy Infrastructure (Proxy Measurements)

Background

As an alternative to measuring CO_2 directly, we can monitor the energy infrastructure of countries to verify whether claimed *actions* are being taken to reduce CO_2 emissions (e.g. construction of wind farms, reduced coal consumption, etc.). Further, we can *infer* the anthropogenic CO_2 emission rate. This is done by monitoring the energy infrastructure, estimating fossil fuel

consumption, and then estimating CO_2 emissions using appropriate emission factors. We refer to these estimates as "proxy measurements." Signatures of energy infrastructure include, for instance, imagery of construction and operation of power plants, oil refineries, and biorefineries, and of fuel transportation activity. Note that monitoring energy infrastructure does not provide information on CO_2 emission other than fossil fuel consumption and does not address greenhouse gases other than CO_2. It can, however, provide information on installed capacity using alternative energy sources.

In the context of international agreements, monitoring of energy infrastructure might include, for example, observations of some of the specific "activities" described in national documentation submitted to the UN Framework Convention on Climate Change (UNFCCC). Proxy observations are useful for informing national policy regardless of the details of the international agreements that are in force at any given time.

Goals of Energy Infrastructure Monitoring

We propose that a useful goal for energy infrastructure monitoring is to be able to characterize the balance between carbon intensive and low-carbon (alternative) energy sources such that the share of the energy from low-carbon sources can be determined with an accuracy of better than ±20% (90% confidence). That is, if the share of the energy from low-carbon sources is actually 25%, the goal of energy infrastructure monitoring should be to determine this share to be 25±5%.

For quantitatively inferring CO_2 emissions from fossil fuel use, we suggest a goal of better than ±20% (90% confidence), i.e. as least as good as suggested for direct measurements.

Assessment of Capabilities: Energy Infrastructure Monitoring

- Monitoring of energy infrastructure is an attractive near-term alternative to direct measurements of GHG concentrations for supporting U.S. monitoring of international agreements. With existing capabilities, im-

portant information can be obtained on a country's use of fossil fuels and alternative energy sources. Consent is not required to implement this approach. The approach can be used to monitor and verify *actions* agreed to by countries as part of mitigation plans or other aspects of international agreements.

- Beyond verifying actions, proxy data from energy infrastructure monitoring can quite plausibly be used to make independent estimates of CO_2 emissions from fossil fuel use. However, it is premature to say what accuracy can be achieved for quantitative estimation of CO_2 emissions using proxy data. Rather, these methods should be assessed by testing their ability to independently estimate the fossil fuel CO_2 emissions from the U.S., or other cooperating country, for which fossil fuel consumption is well-known via other means.

1.4 Roadmap

We summarize the recommendations from our report in the form of a roadmap for action by the U.S. government over the next 2-5 years:

• Acquire and maintain a detailed technical knowledge of the energy infrastructure of countries with large greenhouse gas emissions, and identify and observe the signatures needed to quantify their energy use. This effort requires a significant number of technical personnel and will require standing up of an organization dedicated to this task. This organization should take a systems approach to deciding which signatures and what spatial and temporal sampling will be most useful in quantification of energy use. It should also develop a plan for validating its methodologies against ground truth.

• Fund four strands of development that can be brought to technical maturity within two years:

1. Improved tools and models that provide quantitative estimates of GHG flux error together with an "error budget" that includes measurement

error, uncertainties due to finite sampling density, meteorological, and modeling errors. The tools should also have the capability to support design and optimization of new networks of remote and *in-situ* sensors.

2. Capability to carry out experiments with controlled release of tracer gases to empirically evaluate tools and models.

3. Design-studies of geosynchronous (GEO) satellite concepts brought to enough maturity that capabilities, feasibility, and cost are known. Non-GEO high orbit options should also be investigated as alternatives.

4. A "challenge" exercise to develop a \$500, 1 ppm CO_2 sensor, including symposia to compete designs and trade ideas for sensor design and deployment.

• Following these initial steps, perform systems engineering using modeling/design tools to optimize a "market basket" of investment in point sensors and satellites, assessing tradeoffs and synergy between various possible investments (space vs. *in-situ* sensors, accuracy vs. density of sampling, etc.). Continue the maturation of modeling and design tools.

2 INTRODUCTION

This JASON report assesses methods for the remote determination of greenhouse gases, in particular carbon dioxide. Greenhouse gases (GHGs) are atmospheric gases that are transparent to radiation at one wavelength and absorb at another, usually longer, wavelength. For GHGs, radiant energy can be trapped via absorption, affecting the balance between the radiant energy incident on the earth and that re-radiated back into space. A GHG gas has a *radiative forcing* (measured in units of Watt/m^2/s) and a related *warming potential*. The important anthropogenic GHGs include carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), chlorofluorocarbons(CFCs), perfluorocarbons (PFCs), and hydrofluorocarbons (HFCs).

Figure 2 shows the growth of CO_2 in the atmosphere over time as measured from Mauna Loa. This is the famous "Keeling" curve [1]. The growth of the CO_2 concentration is anthropogenic in nature and primarily due to the burning of fossil fuels.

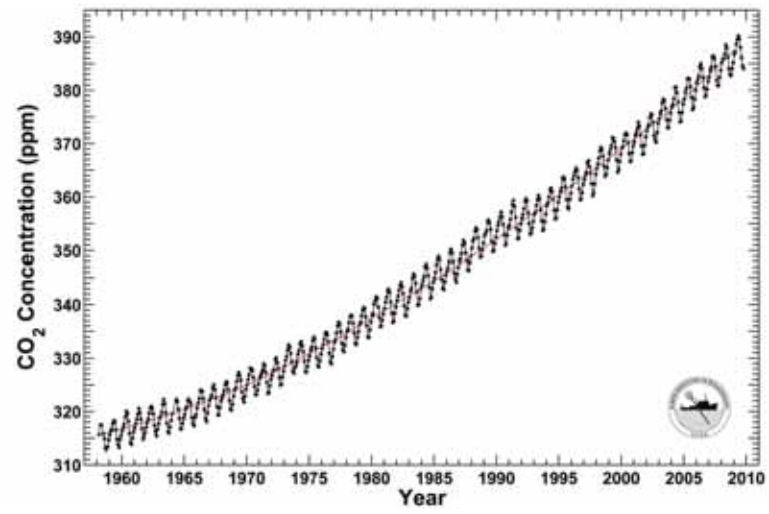


Figure 1: Monthly mean concentration of CO_2 measured at Mauna Loa, Hawaii. SOURCE: Scripps Institute of Oceanography CO_2 program [2]

There are widespread concerns that increased GHG concentrations will have negative consequences for climate and the environment (c.f. [3]), and these concerns have been the impetus for international agreements, including the Kyoto protocols [4] and the more recent Copenhagen accords[5]. Additional international agreements are likely in the future and monitoring of compliance with these agreements will be important. JASON was therefore asked: How might measurements of GHGs be used to support monitoring of international GHG agreements? What are the technical means that can be employed for measurements of GHGs? How might the assessment of such capabilities influence the types of agreements that are entered into? These questions provided the context of our study.

In general, there are three approaches to assessment of greenhouse gas emissions in the context of monitoring of international agreements:

- **Inventories.** Validating a country's self-reported levels of GHG-related activity, in particular national inventories and national mitigation actions.
- **Indirect or "proxy" methods.** Monitoring energy infrastructure and land-use to indirectly infer GHG emissions.
- **Direct measurements.** Estimation of greenhouse gas emissions using direct measurements of GHG concentrations in the atmosphere.

The first two approaches are sometimes called "bottom up" approaches, while direct measurement is considered a "top down" approach.

2.1 Scope of the JASON Study

JASON was asked to assess U.S. capabilities for estimating greenhouse gas (GHG) emissions in support of monitoring international agreements. In particular, JASON was asked:

- What are the current and future capabilities for estimating GHG emissions using *direct measurements* of GHG atmospheric concentrations?
- How do direct measurements compare to *indirect ("proxy") methods* for estimating emissions?
- What are the capabilities for estimating emissions from non-cooperating countries i.e. those for which access to in-country measurements is impeded or denied.

The measurement of GHGs is a large and complex technical area. Our study focused primarily on measurement of the most prominent anthropogenic GHG component, atmospheric CO_2 arising from fossil fuel emissions. The study included the following as primary topics of investigation:

- *Direct measurements* of atmospheric CO_2, including remote observations via satellite and *in situ* measurements using land, air, or ocean-going samplers. The sub-topics included:
  - Modeling, data assimilation, and inversion needed to derive CO_2 fluxes from measured CO_2 concentrations, and to optimize measurement strategies.
  - Isotopic measurements using both stable and unstable isotopes.
- *Energy infrastructure (proxy) data* as an alternative or supplement to direct measurements.

Each of these topics is covered in more detail in chapters of this report.

Our study focused primarily on measurement of *atmospheric CO_2 arising from fossil fuel emissions*. Although we were not able to investigate them in any significant detail, we recognize the importance of several other topics such as: GHGs other than CO_2, GHGs from natural processes and land-use change, and gases other than GHG which correlate with fossil fuel burning.

JASON benefitted significantly from other recent studies on estimation of GHG emissions. The NRC has carried out an excellent study: *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements*, chaired by Stephen Pacala [2]. The study report has three primary chapters: (1) National Inventories of Greenhouse Gas Emissions, (2) Measuring Fluxes from Land-Use Sources and Sinks, and (3) Emissions Estimated from Atmospheric and Oceanic Measurements. It thus covered all three of the basic approaches mentioned above. The NRC report contains a wealth of useful background information as well as useful appendices from which our JASON study benefitted significantly. We also benefitted from the recent study by the MEDEA group.

2.2 Greenhouse Gases, CO_2, and the Carbon Cycle

We focused our study on methods for measuring CO_2 emissions, with particular emphasis on fossil fuel emissions. Figure 2 shows the estimated importance of various GHGs for long-term warming potential. It indicates that CO_2 is the largest contributer, with CH_4 next, then N_2O, followed by about a 1% contribution from other GHGs. Because each GHG has a different lifetime in the global system, the contribution to the warming potential is calculated for a specific time scale, 100 years for Figure 2. At shorter time scales, the relative contributions change; on a 20-year time scale CH_4 has roughly the same importance as CO_2.

For our study we chose to concentrate on CO_2 from fossil fuel combustion, the largest contributor to warming potential on a 100-year timescale. While this was our primary focus, we do recognize the importance of other GHGs, particularly CH_4, and these need to be included in any comprehensive study of direct measurements of GHGs. Additional discussion can be found in [2].

The fossil fuel CO_2 emissions from large countries are often measured in units of billion metric tonnes per year, or Gt/yr. A billion metric tonnes is 10^{15} grams, or a petagram (Pg). Emissions are quoted is some cases for CO_2

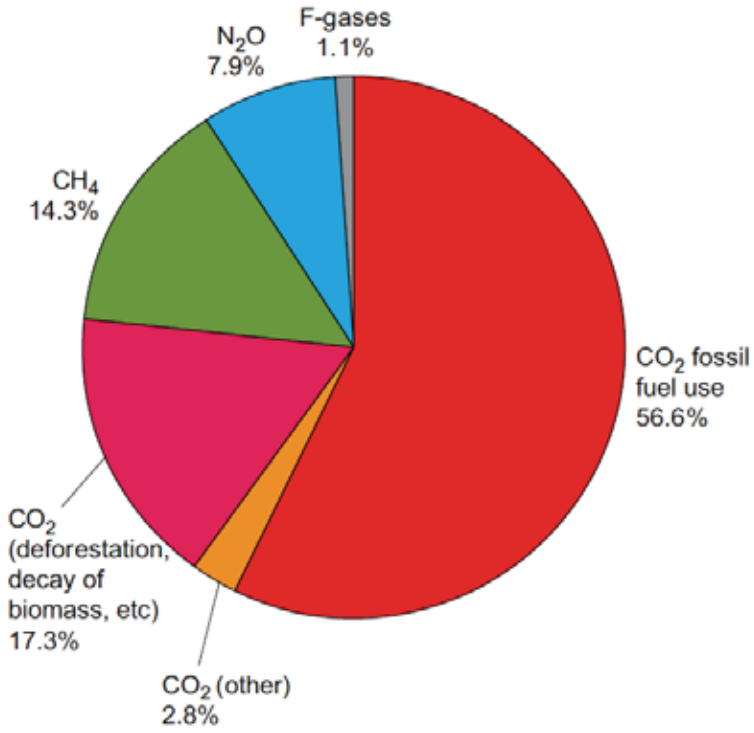


Figure 2: Contribution of individual GHG to the 100-year warming potential. Each GHG has a characteristic lifetime in the atmosphere. Consequently, warming potential is quoted for a specified time in the future. SOURCE: Figure 1.1b from [6]

and in some cases for carbon. Often the notation PgC or GtC is used to indicate carbon emissions, while $PtCO_2$ or $GtCO_2$ indicate CO_2 emissions. The conversion factor is 44/12, the ratio of molecular weight of CO_2 to the atomic weight of the most abundant isotope of carbon.

Although CO_2 from fossil fuel burning represents the major contributor to the long-term GHG warming potential, it is only a small component of the overall exchange of CO_2 between the atmosphere and the land/ocean reservoirs. Figure 3 indicates that the global burning of fossil fuels released about 6 Petagrams (Pg) of carbon per year into the atmosphere in the timeframe of the year 2000. This is to be compared to the much larger exchange between the ocean and atmosphere of roughly 90 Pg/yr and the exchange of CO_2

between the terrestrial biosphere and the atmosphere of about 120 Pg/yr. Each of these exchanges has significant natural variability on a wide variety of both spatial and temporal scales. Thus any attempt to estimate anthropogenic CO_2 emissions using direct measurements must account for the significant variability in the natural carbon cycle. This is discussed in more detail in Section 4 which discusses direct measurements of CO_2. Note also that because the ocean uptake is thought to be larger than the ocean loss, the growth of CO_2 in the atmosphere is less than the input due to anthropogenic sources. A more complete discussion of the global carbon cycle and natural sources and sinks of CO_2 can be found in [2] and references therein.

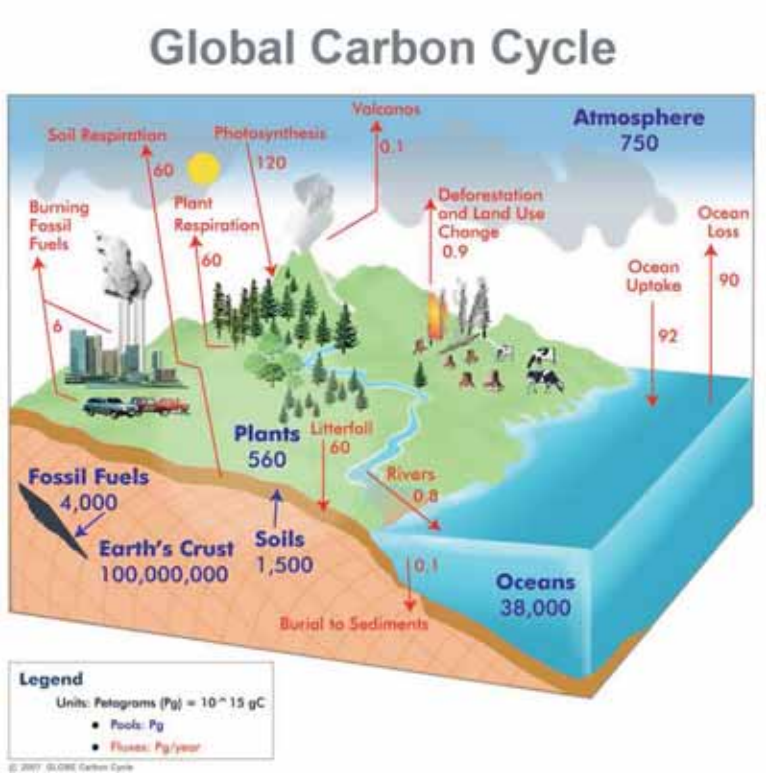


Figure 3: The global carbon cycle, showing reservoirs (pools) and fluxes. SOURCE: The Globe Program, Carbon Cycle Project.

2.3 Estimating CO_2 Emissions from Fossil Fuel Burning

Existing international agreements (e.g. [4], [5]) are often stated in terms of reductions in CO_2 emissions. Monitoring of international agreements therefore requires estimates of anthropogenic CO_2 emissions (or fluxes) from specific countries or regions. However, the relevant quantities that can be measured directly are not CO_2 emissions but rather CO_2 concentrations. Specifically, these include: (1) *in situ* measurements of CO_2 concentrations (mole fraction in parts per million) obtained by sampling, and (2) CO_2 column concentrations (weighted mole fraction) looking down at a patch of the Earth from an orbiting satellite. To use measurements of CO_2 concentrations to obtain an estimate of CO_2 emissions requires a model for the transport of CO_2 from the source to the measurement point. In particular accurate meteorology is required, i.e an accurate measurement or model of wind velocity, temperature, pressure, and other variables. Thus the performance of a measurement system for CO_2 emissions (fluxes) depends both on the concentration measurements themselves and on what is called the "model inversion" to go from the measured concentrations to the derived emissions. We cannot overemphasize the importance of transport modeling in estimation of CO_2 emissions. In many cases, modeling uncertainties will dominate measurement uncertainty. We discuss modeling, data assimilation, and inversion further in Section 5.

As mentioned before, the naturally occurring carbon cycle includes CO_2 exchange to and from the atmosphere by processes in the biosphere, soils, and oceans, as well as geophysical processes (e.g. volcanos). Assessment of emissions due to human activities is therefore difficult because the CO_2 emitted by anthropogenic sources is only a small fraction of the natural emission and absorption processes.

It is useful to estimate the size of the average enhancement in concentration downwind from an anthropogenic source of CO_2. A simplified estimate of the downstream concentration enhancement is given by (c.f. [2]):

$$
\begin{aligned}
\Delta &= \frac{F_{emit}}{A} \times T_{dwell} \times \frac{1}{\rho_{col}} \\
&= 4 \text{ ppm} \left(\frac{F_{emit}}{1 \text{ Pg/yr}} \right) \times \left(\frac{A}{10^7 \text{ km}^2} \right)^{-1} \times \left(\frac{L_{path}}{3 \times 10^3 \text{ km}} \right) \\
&\quad \times \left(\frac{v_{wind}}{5 \text{ m/s}} \right)^{-1} \times \left(\frac{\rho_{col}}{4 \times 10^4 \text{ mol/m}^2} \right)^{-1} \\
&= 0.4 \text{ ppm} \left(\frac{F_{emit}}{1 \text{ Pg/yr}} \right) \times \left(\frac{A}{10^7 \text{ km}^2} \right)^{-1} \times \left(\frac{L_{path}}{3 \times 10^3 \text{ km}} \right) \\
&\quad \times \left(\frac{v_{wind}}{5 \text{ m/s}} \right)^{-1} \times \left(\frac{\rho_{col}}{3.6 \times 10^5 \text{ mol/m}^2} \right)^{-1}
\end{aligned}
\tag{2-1}
$$

where Δ is typically given in parts per million (ppm) and F_{emit} is the total emission from a region of interest having an area A. The dwell time of the measured air in the emitting region is $T_{dwell} = L_{path}/v_{wind}$ where L_{path} is the path length across the emitting region and v_{wind} is the average wind speed. Note that (L_{path}/A) characterizes the fraction of the total emission, F_{emit}, that is accessible to the measurement. The total atmospheric column density, ρ_{col}, is the column density appropriate to the mixing volume or the measurement column. For an *in situ* measurement the column is typically that of the mixed boundary layer i.e. of order 1 km of the lower atmosphere. For satellite measurements, the column is the total column depth of the atmosphere.

As can be seen, a typical enhancement is 4 ppm in the boundary layer and 0.4 ppm for a satellite measurement of the enhancement compared to the total atmospheric column density. Additional estimates of enhancements are given in the next section. These should be compared to the total CO_2 dry mole fraction of about 400 ppm which can have substantial diurnal, synoptic, and seasonal variability, much larger than the 0.4 ppm (satellite) and 4 ppm (*in situ*) enhancements from anthropogenic sources.

2.3.1 Country and regional fossil-fuel CO_2 emissions

Relatively few countries/regions contribute the majority of the CO_2 emissions from fossil fuel burning. Graphically the situation is shown in Figure 4 for the year 2004 and in table format in Table 1 for the year 2008. Including estimated fossil fuel emissions from the United States, China, and the European Union accounts for approximately 55% of CO_2 emissions. Adding Russia, India and Japan brings the total to approximately 70% of fossil fuel CO_2 emissions. In this report we therefore suggest concentrating initially on estimates of emissions from the United States, China, and the European Union, and secondarily from Russia, India, and Japan.

Table 1: Carbon Emissions from Energy Consumption. Source: Adopted from US EIA estimates of CO_2 emissions. Note that the units are petagrams of carbon, rather than petagrams of CO_2. (For comparison, the EIA estimated the world emissions in 2000 to be 6.5 petagrams, or about 24 petagrams of CO_2.)

Carbon Emissions from Energy Consumption		
Region	2008 Emissions (Petagram Carbon)	(%)
World	8.15	100
China	1.70	21
United States	1.64	19
EU-27	1.16	14
Russia	0.45	5.7
India	0.38	4.9
Japan	0.34	4.0

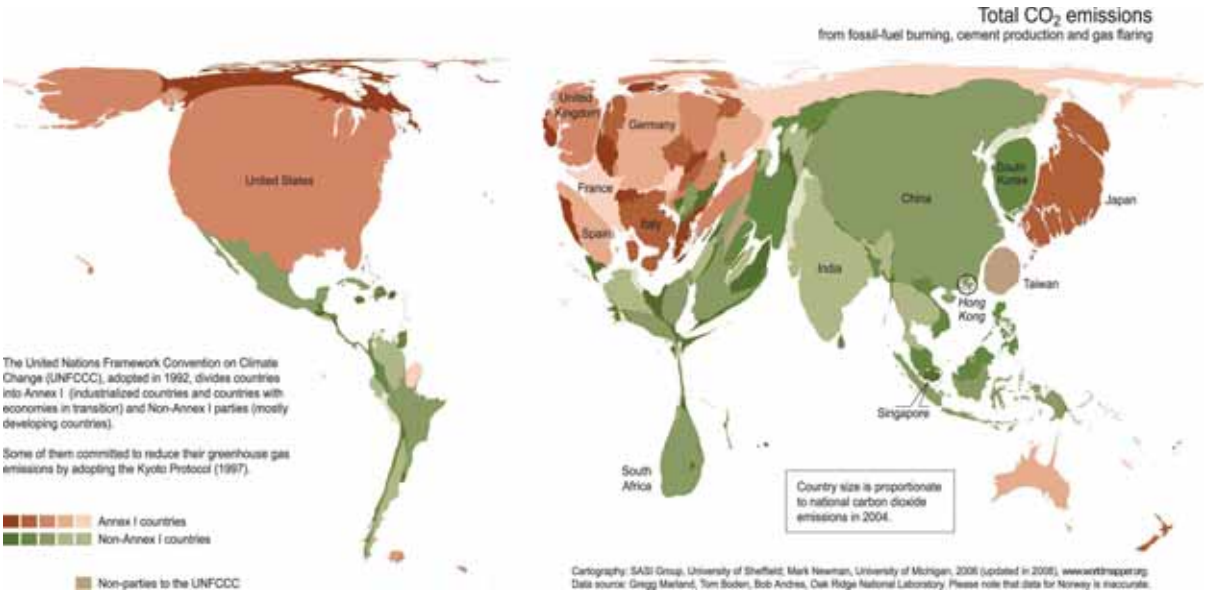


Figure 4: Total CO_2 emissions. Country size is proportional to national CO_2 emissions in 2004. SOURCE: Oak Ridge National Laboratory.

Table 2: Average Downwind Enhancements for Selected Cities and Countries (2005 Estimates). SOURCE: [2].

Country/City	Area (km^2)	CO_2 Emissions ($GtCO_2$/yr)	Total Column (ppm)	Boundary Layer for 1 km (ppm)
United States	9,827,000	5.9	0.76	6.8
China	9,571,000	5.6	0.73	6.5
Russia	17,075,000	1.57	0.15	1.4
Japan	378,000	1.25	0.82	7.3
India	3,166,000	1.22	0.28	2.5
Los Angeles	3,700	0.073	0.49	4.3
Tokyo	1,700	0.064	0.63	5.6
Beijing	800	0.074	1.1	9.4

For reference, in Table 2 we provide some typical downstream enhancements using Eq. 2-1 for the US, China, and for several large cities as given in the recent NRC report [2]. We also note that the enhancement 2 km downwind from a 1 GigaWatt coal-fired power plant will be about 2 ppm for a total column measurement and about 15-20 ppm for an *in situ* measurement in the boundary layer.

2.3.2 Benchmarks for accuracy of CO_2 emission estimates

NOAA's CarbonTracker Project (see [7]) provides an example of the accuracy currently obtained for estimates of CO_2 emissions using an *in situ* sampling system. There are CarbonTracker projects in both the US and Europe, with the European effort carried out in collaboration with Wagenigen University. Figure 5 shows the CarbonTracker Europe sampling network and Figure 6 shows emissions estimates obtained for Europe between 2001 and 2007. It is important to recognize that the fossil fuel estimates shown in Figure 6 are not direct measurements, but rather bottom-up estimates. These plots therefore represent a direct measurement estimate of the yearly biospheric and ocean fluxes assuming that the bottom-up fossil fuel estimate is correct. The emphasis of CarbonTracker to date has been on estimation of biosphere uptake and loss, not on an independent direct estimate of fossil fuel emissions. However, CarbonTracker does hope to introduce a process model for fossil-fuel CO_2 production in the future and then perform a joint optimization and obtain estimates of both of the principal components, the biosphere and fossil fuel burning. It is plausible that the errors on the direct estimates of fossil fuel CO_2 emissions will be comparable to the errors on the biospheric component, i.e roughly 30% of the net flux, or about 0.5 PgC/yr, 1-sigma. Combining *in situ* estimates with satellite observations should further improve the achievable accuracy for anthropogenic CO_2 emissions.

It is useful to set a benchmark to assess future progress in estimation of anthropogenic CO_2 emissions for the purpose of monitoring international agreements. Ideally, we would like to set a goal of 10% or even better accu-

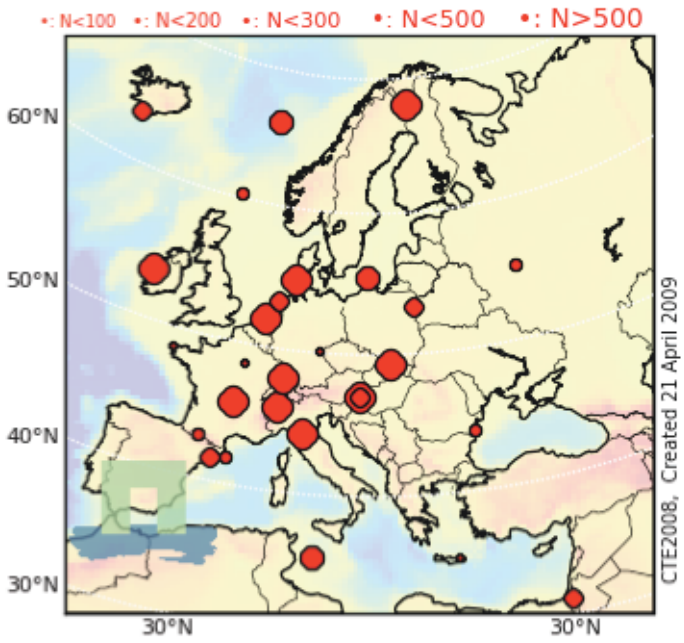


Figure 5: *In situ* sampling network used by CarbonTracker Europe. SOURCE: CarbonTracker Europe

racy for the measurement of anthropogenic CO_2 emissions from a major GHG emitting country. However, given the current demonstrated state-of-the-art in direct measurements of CO_2 and given the uncertainties in transport modeling we prefer to set a less stringent benchmark as an initial step towards achieving useful direct measurement capability for CO_2. JASON has therefore chosen $\pm 20\%$ (90% confidence) measurement uncertainty as an initial benchmark by which to evaluate the ability of the US to estimate annual anthropogenic CO_2 emission from a country or region. This benchmark was chosen as a compromise between what we believe is achievable and what is required for treaty monitoring. It assumes a modest level of reasonable prior information, which is combined with the direct measurements to yield an emissions estimate. This prior information might include: process models for biospheric processes, process models for anthropogenic sources of CO_2, and population density information. Such prior information is typically required

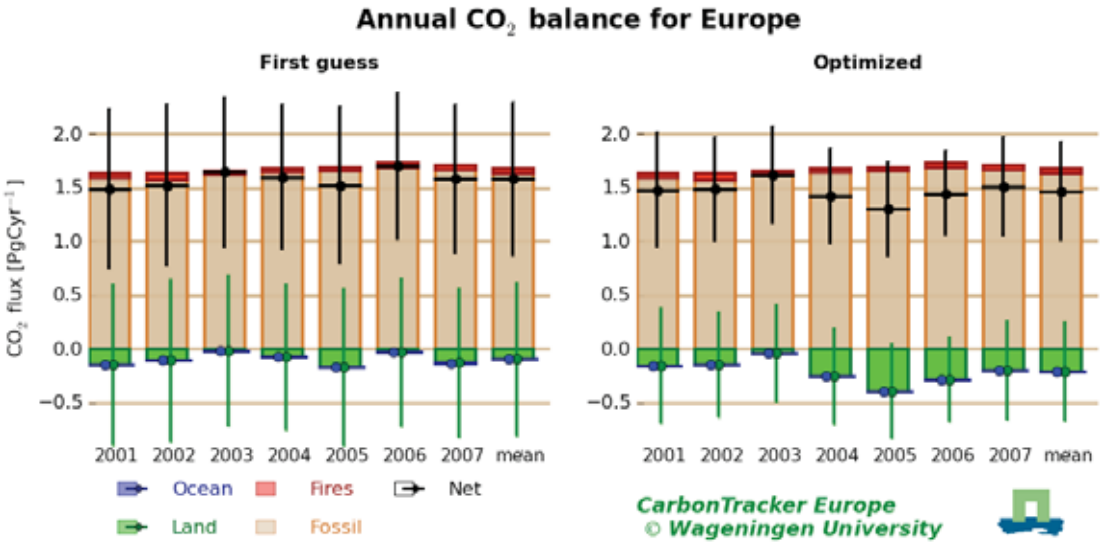


Figure 6: Emission estimates for net flux and terrestrial biospheric flux derived from CarbonTracker Europe. The left hand panel is an initial guess at the fluxes using land and ocean process models. Fossil fuel and fire estimates come from a "bottom-up" estimate. The right hand panel shows the estimates after incoporation of data from the *in situ* sampling network. Note in particular the errors in the right hand panel which are about 0.5 PgC/yr or about 30% of the estimated net flux. See text for discussion. SOURCE: CarbonTracker Europe

to derive estimates from a limited sample of concentration measurements (see Section 5).

As discussed above, the largest GHG emitting nations include: China, the U.S., Russia, Japan, and India, as well as the European Union as a region. For some of these, direct measurements at the ±20% level will be difficult to achieve as it is not only the total emission level that is important, but also the emission density (i.e. tonnes of CO_2 per square kilometer per year, see Table 2). We suggest starting with the larger emitters: the European Union, China, and the U.S. with the goal of developing a direct measurement capability for these at the ±20% level or better. We are optimistic that these levels can be achieved through a combination of satellite and *in situ*

measurements. Once this capability has been demonstrated, the accuracy of both the measurements and the transport modeling will be better understood and it should be possible to better ascertain whether a further improvement to $\pm 10\%$ accuracy, or better, is readily achievable. The ultimate technical goal should be to improve the satellite and *in situ* sampling networks to the level that estimates of CO_2 emissions are limited by natural variability rather than by measurement and modeling errors.

2.4 Monitoring of Energy Infrastructure as a Near-Term Alternative to Direct Measurements

Direct measurements of GHGs are critically important. The goal of international agreements on GHGs is to limit the effect of the growth of GHGs on climate and the environment. Consequently, direct measurements of the gases are needed to validate the ultimate efficacy of the treaties in limiting GHG growth. While bottom-up approaches such as inventories and proxy data are needed to help monitor compliance with international agreements, it is only through measurement of the gases themselves that it can be demonstrated that the agreements are having their desired effect of reducing GHG production. This is particularly true for CO_2 from land-use and GHGs other than CO_2 which are more difficult to monitor accurately through bottom-up approaches.

However, as discussed above, direct measurents of anthropogenic GHGs are very challenging, particularly for CO_2 for which the signals are small, variability is high, and transport modeling difficult. Progress in improving the accuracy of direct measurements will likely take place on a 5-10 year timescale. As a consequence, it is useful to consider indirect bottom-up approaches as a near-term alternative to direct measurements.

In Section 3 we discuss monitoring of energy infrastructure as a useful component for monitoring international agreements. Characterizing and quantifying the enegy infrastructure of large GHG emitting countries has two ben-

eficial aspects. First, it can monitor compliance with promised *actions*, for instance, transition from a carbon intensive to a low-carbon energy system. In particular, monitoring of the energy infrastructure can verify installed capacity for production of energy from non-fossil energy sources, something that direct measurement of GHGs will not do. Second, accurate knowledge of the energy infrastructure of a country could yield a "proxy" estimate of CO_2 emissions by assuming emission factors for various supply and demand sectors a country's energy system. These proxy estimates will be useful for comparison with direct measurements. Further, knowledge of the energy infrastructure will allow better process models to be built for CO_2 emissions, which will be useful in the transport modeling required for direct measurement approaches.

An attractive aspect of energy infrastructure monitoring is that the technical basis for such monitoring already exists and an operational system can be deployed quickly. This can provide a near-term capability for monitoring of actions promised under international agreements. We emphasize the difference between monitoring *actions*, which can be done in the near term, and *estimating CO_2 emissions* by characterizing energy infrastructure, which will have a distinct set of challenges and take longer to develop into an operational capability.

We discuss monitoring of energy infrastructure in the next section of this report.

3 ENERGY INFRASTRUCTURE MONITORING

Much of this study concentrates on development of point and remote sensors that can be used to deduce the CO_2 emissions of a country and/or region of interest, with or without auxiliary information about deployed energy infrastructure. In this respect, the sensing problem is being approached neutrally, without regard to the details of the transition in energy infrastructure that would be required to implement a low-carbon energy system on the scale of a country or a region. In the following we discuss the telltale signatures of a low-carbon energy system, how they could be monitored, how such signatures can be used to evaluate whether a country has actually installed low-carbon energy systems or not, and how to assess the extent of continued use of fossil fuels.

Note that in this process we do not necessarily seek to establish quantitatively whether a country has reduced emissions by 80%, or 70%, or 90% relative to some baseline level, although this may be a secondary objective. Rather, the first objective of energy infrastructure monitoring is to confirm that *actions* are being taken to install generation capability from alternative energy sources (e.g. construction of wind farms, etc.) and that *actions* are being taken to reduce generation by fossil fuel sources (e.g. reduction in oil imports or refinery throughput).

We also note that the approaches outlined here are primarily useful for characterizing actions related to reductions in anthropogenic fossil fuel CO_2 emission, and are not as useful for deducing the net carbon dioxide emissions of a region, nor are they as useful for GHGs other than CO_2.

Low-carbon energy systems are characterized by only a limited set of choices for electricity generation and for fuel consumption. For example, if emissions are to be eventually cut by 80%, then essentially there can be no operation of natural gas utilities or associated natural gas infrastructure that supplies

individual consumers, e.g. for heating of buildings. Otherwise, the emissions associated with such natural gas consumption alone will exceed the region's total CO_2 emissions budget. Similarly, all electricity generation must be essentially carbon-free, meaning either fossil fuel combustion with full carbon capture and storage, nuclear power, and/or large-scale renewable power, primarily from wind and solar electricity generation. Crude oil consumption needs to be greatly reduced, being reserved for use by aircraft, ships, and heavy-duty vehicles, or these parts of the transportation sector need to be served by biofuels, with the remaining oil consumption devoted to serving other parts of the transportation sector, such as in light duty vehicles including plug-in hybrids for example.

Hence, one approach to verifying whether a country/region is reducing emissions would involve means of estimating the amount of low/no carbon energy generated and comparing this value to what would be needed to supply the countrys demand in a low/no carbon fashion. In this approach, we are describing methods to verify the installation and deployment of tomorrow's energy system, instead of verifying incremental or modest emissions reductions in today's carbon-intensive energy systems.

We emphasize several points with regard to this approach:

- The emphasis is on monitoring actions that accompany energy infrastructure, rather than monitoring or estimating net CO_2 emissions.
- The methodology can be used to monitor compliance with international agreements to undertake actionable items, e.g. increasing the share of energy generated by non-fossil fuels.
- Technical methods currently exist that can be used to monitor energy infrastructure of large GHG emitting countries. They benefit from, but do not require, the cooperation of the emitting country.
- The large changes in the energy infrastructure during a transition to a low-carbon system have large signatures that can be monitored using

many different methods. Extensive detailed information already exists on the energy infrastructures of large countries. This can be supplemented with visible and thermal remote sensing to validate and extend such information.

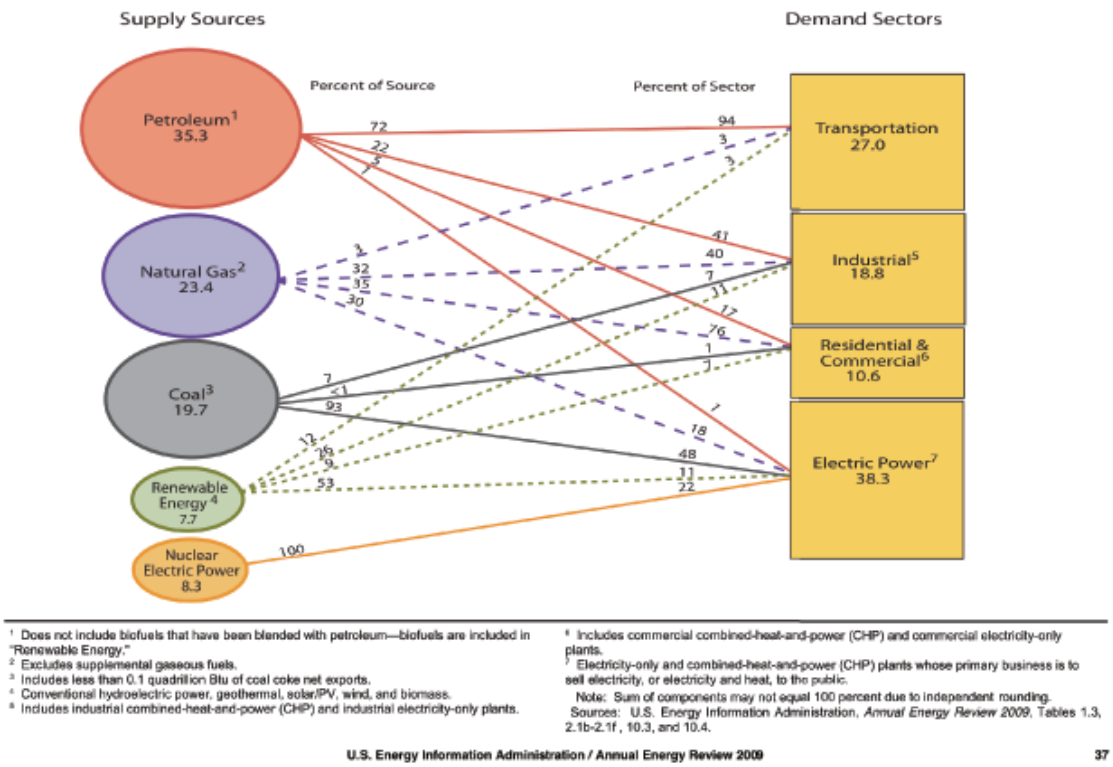
- A similar methodology can be applied to monitoring actions related with carbon offsets.
- It is important to begin immediately to establish baselines for the energy infrastructures of large GHG emitting countries.

3.1 Monitoring Supply Sectors

A system for monitoring energy infrastructure can either focus on monitoring the supply sector of the energy infrastructure, or the demand sectors. Figure 7 is a recent estimate by the DOE Energy Information Administration of the energy flow in the US energy system. To understand the transition from a carbon-intensive system to a low-carbon transition, the relative sizes of the fossil-fuel and non-fossil fuel supply sources need to be estimated. The approach we suggest is to monitor the "choke points" in the energy supply system i.e. quantitatively assessing the size of various sectors: the electricity generation sector, the natural gas distribution sector, the coal production and transportation sector, and the oil transport and refining sector. We discuss each of these these in more detail in following sections.

We envision a monitoring system that collects input from multiple sources: visible and thermal imaging satellites (both civil and commercial), available data on energy production and consumption (compiled by both private and governmental organizations), and industrial data on the characteristics of key components of an energy infrastructure (e.g. refinery design, electric turbine characteristics, nuclear power plant operating characteristics, etc.).

Figure 2.0 Primary Energy Flow by Source and Sector, 2009
(Quadrillion Btu)



[1] Does not include biofuels that have been blended with petroleum—biofuels are included in "Renewable Energy."
[2] Excludes supplemental gaseous fuels.
[3] Includes less than 0.1 quadrillion Btu of coal coke net exports.
[4] Conventional hydroelectric power, geothermal, solar/PV, wind, and biomass.
[5] Includes industrial combined-heat-and-power (CHP) and industrial electricity-only plants.
[6] Includes commercial combined-heat-and-power (CHP) and commercial electricity-only plants.
[7] Electricity-only and combined-heat-and-power (CHP) plants whose primary business is to sell electricity, or electricity and heat, to the public.
Note: Sum of components may not equal 100 percent due to independent rounding.
Sources: U.S. Energy Information Administration, *Annual Energy Review 2009*, Tables 1.3, 2.1b-2.1f, 10.3, and 10.4.

U.S. Energy Information Administration / Annual Energy Review 2009 37

Figure 7: US Energy Flow by Source and Sector. SOURCE: DOE/EIA

3.1.1 Electricity generation sector

a) Nuclear power

The generation capacity of a nuclear power plant can be readily established by its design parameters, as well as by the amount of fuel consumed. Capacity factors for nuclear power plants are typically 90%, and it is reasonable to assume that no one is going to build a large nuclear power plant and declare unused capacity merely for the purpose of carbon treaty compliance (the cost of not running the plant at 90% of its rated capacity and bearing the cost of that underutilized capital, which is a large fraction of the total cost of nuclear

electricity, is just too high). Optical imaging with moderate resolution, such as that available on Google Maps, can readily reveal the number and location of nuclear power plants in a country or region. Thermal imaging might also be useful in estimating the amount of power being generated at a particular time by each power plant; the electrical power output of a given reactor design (such as the Westinghouse AP1000, for example) is well known and it might thus be possible to to roughly constrain the output generation power at the time that the facility is being interrogated.

b) Solar power

The average insolation of any region of a country is well-known, and is publicly available on readily accessible maps. Additionally, the capacity factors of both solar thermal and solar photovoltaic systems are also well-known. Solar photovoltaic capacity factors are approximately 0.2, averaged over a year; relative to peak capacity, a factor of 2 reduction occurs due to the diurnal cycle, another factor of 2 reduction occurs due to time-of-day variation in insolation for non-tracking, flat plate, systems, and the remaining reduction results from the averaged level of obscuration due to cloud cover. This last factor is also well known for different regions as well, and so local corrections on a yearly averaged basis can be readily applied to a given system.

A survey of the total area of installed solar panels, their angle of inclination, and their location can thus yield an estimate of the total electrical energy generated by that system in a year. Similarly, a survey of the number of concentrated solar thermal dishes, troughs, etc., along with the local insolation and capacity factors, can be used to estimate the power production from utility-scale solar thermal installations.

Note that to make a material contribution to electrical energy production, very large areas, and very large numbers of, such solar systems will need to be installed. They also do not go away and hence even an occasional, annual or less often, overflight, will suffice to reveal the extent of such installations in the region of concern.

c) Wind power

The wind energy potential is well-known for most of the world, both at 50 m hub heights and at 80 m hub heights, and can be found in maps that are readily available on open sources. Capacity factors of wind turbines are also well-known, and depend primarily on the class of the local wind resource. The peak capacity of a wind turbine is readily estimated from the physical parameters of the turbine blade. Typical modern wind turbine blades are 50 m in length, with next generation blades being perhaps 80 m in length. Extensive generation of electricity from wind will require installation of a very large number of wind turbines. Information on the dimension of the turbine blades could be obtained from optical imaging. Additionally, blade dimensions, as well as the rotation speed of wind turbines in the wind farm, can be obtained through use of radar information that is now considered an annoying issue of concern with respect to interfering with air traffic control radars (c.f. [17]).

d) Hydroelectricity

Hydropower is essentially maxed out globally, so it is not really necessary to monitor hydroelectric generation for expanded deployment associated with additional electricity production that would be needed by a country to transition to a low-carbon energy system. Nevertheless, it is possible to estimate the amount of hydropower that a country or region is producing. The surface area, and volume, of most major reservoirs is well-known and can readily be found in open sources. In addition, the amount of hydroelectricity produced in a year from a specific reservoir can be estimated by using satellite altimetry to monitor the change in height of the water in the reservoir at selected times, presumably at the end of winter and at the end of the summer, for example. TOPEX/Poseidon can determine sea level with a precision of 4.6 cm, for example. Knowledge of the head height of the dam and the fact that hydroelectric power generation is typically 85-90% efficient, then allows facile calculation of the amount of hydroelectricity produced by a given reservoir in a year. For example, see [13]. Summation over an entire region and/or coun-

try thus yields the amount of hydropower produced by a country of interest in a given year.

3.1.2 Electricity distribution system

Essentially all low-carbon energy systems have a significant increase in electricity generation as a percentage of delivered energy relative to todays fossil-dominated infrastructure. In addition to monitoring the generation sources (either distributed, as in the case of wind and solar electricity, or localized, as in utility-scale fossil-fired power plants or nuclear power plants), the supply of the electrical system might be further monitored by monitoring key loss points in the system. For example, step-down high-voltage transformers have a 30 year lifetime (or more), and all exhibit inductive losses on the order of 1-2%. These losses produce a thermal signature that could be monitored, to estimate the level of power supplied along key distribution nodes in the electrical grid. Thermography from aerial overflights is used, for example, routinely, for monitoring and inspection of transmission lines, joints, and other parts of the electrical transmission grid (c.f. [14]). The quantitative accuracy of such remote estimates of energy usage could be assessed by making observations of electrical distribution systems in the U.S. and compared to actual usage data.

3.1.3 Fossil fuel production, distribution, and sequestration

a) Oil

For a country that has little indigenous oil production capability and/or resources, monitoring oil imports is relatively straightforward: oil is supplied on super-tankers, whose capacity is well-documented (and which are also readily monitored by a variety of means). The quantity of oil delivered is also readily available from commercial transactions on the oil market. To first order, all of this oil is burned, so the associated carbon emissions are readily calculated.

For a country such as China or Brazil, that has significant domestic production (for China, total oil consumption is 7.8 million bbl/day, of which 4.0 million bbl/day is produced domestically and 3.9 million bbl/day is imported [9]), one needs an independent method to estimate the amount of oil consumed by the country/region. One approach is to monitor the country's oil refineries. All refineries have prominent distillation towers: the first step in processing of crude is to distill the crude, so as to separate the crude oil into various cuts, depending on their boiling point. The capacity of refineries around the world is well-known and documented in open sources [10], such as the Energy Information Agency (EIA). Refineries typically run at 80-85% of capacity, averaged over the year. The thermal signature of the distillation tower can be measured for existing refineries and then used to roughly gauge the throughput of that refinery in the future as well. Although the specific heat of crudes can be different, they are not different by very much between crude types.

There are also opportunities to reconcile estimated crude oil refining in the supply side with estimates of petroleum usage on the demand side, at least in the late stages of transition from a carbon-intensive to a low-carbon energy system. In a low-carbon economy, essentially all liquid hydrocarbon-based fuels will have to be used for aircraft, ships, and heavy-duty transportation, for which there is no currently credible alternative to liquid hydrocarbon fueling. Collectively, these markets comprise approximately 40% globally of the transportation sector demand for fuel; hence if emissions are to be cut by 80%, light duty vehicles can not run on hydrocarbon fuels obtained from refining crude oil. Hence verification of compliance would involve verification that essentially all emissions from mobile, light duty vehicles are not derived from liquid hydrocarbons produced by refining of crude oil. To do this, we do not propose to monitor the tailpipes of a large number of light-duty vehicles. A useful approach might involve spot collection of samples that can be subjected to analysis for their $^{14}C/^{12}C$ content. The ^{14}C content of the CO_2 in the troposphere spiked in the 1960s as a result of nuclear testing. Emissions of CO_2 derived from fossil fuel combustion (which has

essentially no ^{14}C content) reduces the ^{14}C content of the air in the region of combustion. Especially prominent reductions in ^{14}C have been observed in major cities, due to emissions from vehicles, and around fossil-based electric power plants. It should suffice to verify a reduction in emissions arising from transportation in selected areas such as major cities and the sampling need only occur on an annual basis. Hence a tree can be analyzed for its $^{14}C/^{12}C$ content through a series of tree rings; leaves on deciduous trees that drop in the autumn can be analyzed to yield the $^{14}C/^{12}C$ content for a locality in a given year, or embassies can grow an annual potted plant and send a stem or leaf for analysis, to obtain the desired information, which is collected using accelerator-based mass spectrometry. In a low-carbon energy economy, there should be little depletion of ^{14}C in these air samples no matter where they are collected. See Section 4 and Appendix D for further discussion of how isotopic measurements can be used to characterize the processes that helped produce the CO_2 in a given sample.

b) Gas

A low-carbon economy can have essentially no distributed natural gas emissions. Hence natural gas pipelines to places other than utility electrical generation facilities equipped with CCS should not be in operation. Compliance can be ascertained as follows: a) all natural gas pipelines leak, and the fugitive gas leaks can be detected using remote sensing methods in an operating pipeline but should be absent in a pipeline whose usage has been discontinued; b) pumping of gas through a pipeline requires pressurization of the gas to overcome the frictional dissipative losses of the gas in the pipeline. The pumps will dissipate heat, and when operative, will produce a thermal signature that can be detected by remote sensing techniques.

Compressors can be of various sizes, typically ranging from 75 kW to 2000 kW in size. The dimensions of most installed natural gas pipelines are well-known in the open literature, and dimensions of any newly constructed pipelines would be readily ascertained during the construction phase. The pumps are sized to overcome the frictional dissipative losses associated with gas moving

at a certain velocity through the pipeline; these losses go as the cube of the gas velocity. Compressor stations are typically 15-22 acres in size and are placed at 40-100 mile intervals along the pipelines. Locations of major pipelines and compressor stations are generally well known.

c) Coal

The Daqin line in China carries 1.2 million tons of coal per day. The Power River Basin in Montana/Wyoming produces over 1 million tons of coal per day as well. These open top rail cars are generally maximally loaded by weight with coal, of widely known rank. These rail shipments are well over a mile long and would be hard to miss in optical imagery. A statistical sample of coal rail shipments achieved through imagery would yield useful information on coal consumption patterns.

These rail cars have to go somewhere to deliver their coal. As an example, a map of coal reserves in China is displayed in Figure 8. This coal will either will go to fixed sites, for electrical power generation at utilities, or to dispersed sites, such as to industries. Currently in China about 50% of the coal is used for electricity production at utilities, and the remainder goes to industry. In a low-carbon energy system, essentially no coal can be combusted at locations other than fixed sites which are equipped with carbon capture and storage systems. One can assume that any coal headed to sites that are not equipped with CCS capability produces a stoichiometric amount of CO_2 released to the atmosphere upon combustion. The rank of the coal is well-known based on its origin, and hence the emissions associated with combustion of this coal can be readily estimated as well.

Major coal-burning sites that are equipped with CCS will have a variety of observable signatures. Of course, the number of coal cars that are entering the site can be estimated using the statistics of observations of the coal trains. In addition, the facility will have a detectable and characteristic thermal signature that will be associated with its power production capacity and with its operational protocols (see e.g. Figure 9). The economic sweet

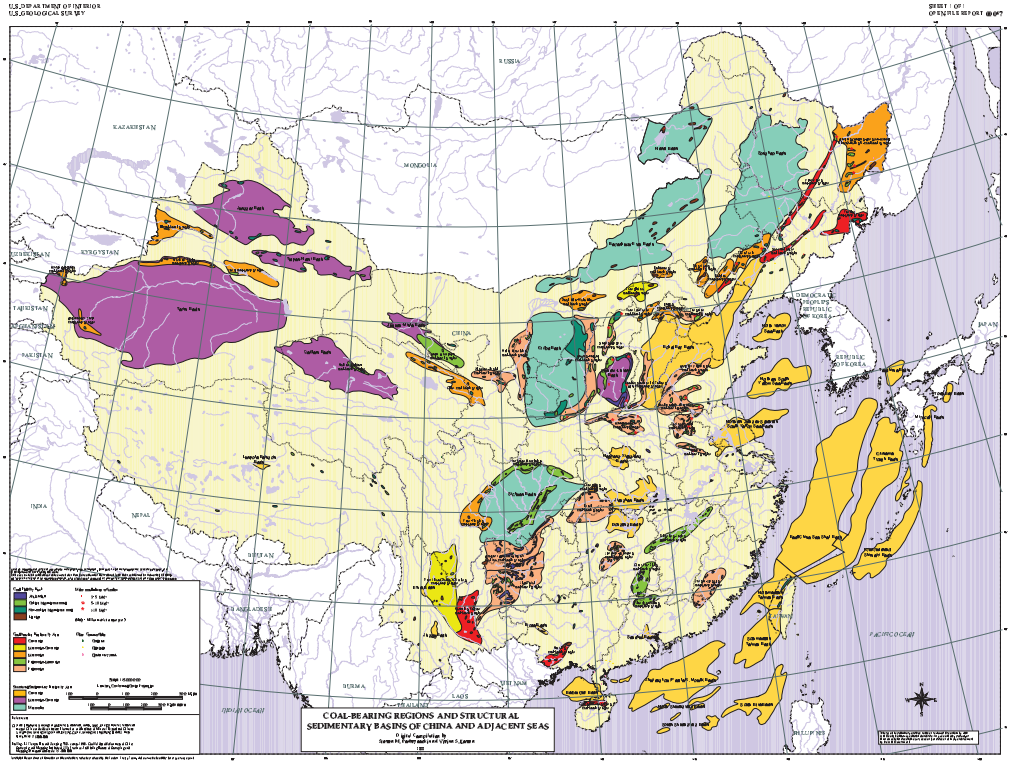


Figure 8: USGS map of China Coal fields. SOURCE: USGS.

spot for CCS is about 80% capture of CO_2 in the flue-gas, when the only currently proven method, involving the use of MEA (methyl-ethanol-amine) as the capture medium, is employed. Hence a detectable concentration of CO_2 will be present in effluent from the CCS facility. The desorption reactor, in which CO_2 is released from the MEA, will have a thermal signature. The compressors that are needed to pressurize the CO_2 for pumping the gas underground will have significant thermal signatures (c.f. [8]), typically being sized at 60 MW for a 500 MW advanced pulverized coal plant, and 30 MW for a 600 MW integrated gasification combined cycle (IGCC) plant. Calibration of these thermal signatures and CO_2 effluent signatures would likely suffice to yield a fairly robust estimate of the power produced by the facility as well as the amount of CO_2 directed under the ground and also of the amount that is escaping out of the facility in the effluent gas stream.

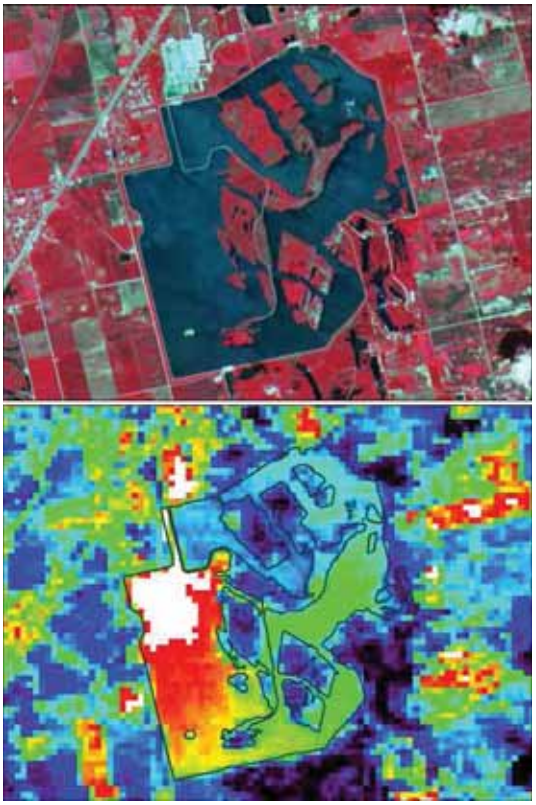


Figure 9: ASTER satellite false-color images acquired over Joliet 29, a coal-burning power plant in Illinois. Joliet 29 can be seen in the VNIR image (top) as the bright blue-white pixels just above the large cooling pond. In the bottom image, a single ASTER Thermal Infrared band was color coded to represent heat emitted from the surface. The warmest areas are white; progressively cooler areas are colored red, orange, yellow, green, blue, and black for the coolest. Note the bright white plume of hot water discharged from the power plant. SOURCE: http://visibleearth.nasa.gov/view_rec.php?id=1716

3.2 Monitoring Demand Sectors

Our study investigated ways to monitor a country's energy infrastructure using data and observations of energy *supply sources*. It seems to us that this is the most straightforward and robust approach to characterizing the large scale energy infrastructure. However, this approach could also be complemented with data and observations from the *demand sectors*. The estimates from both approaches could be compared and checked for consistency.

An example of a project which focuses on the demand side is the Vulcan Project [16]. The Vulcan Project has produced maps of US energy consumption with higher spatial/temporal resolution (10km x 10km) than that provided by typical inventory analyses (1 deg x 1 deg) [11]. Higher spatial resolution requires detailed data on many aspects of energy consumption: detailed road/highway usage statistics, population and building characteristics, etc. Gurney et al. [11] show that the Vulcan estimates yield results very similar to the DOE/EIA inventory estimates, but that their spatially resolved CO_2 emission estimates are significantly different than earlier estimates calculated on a 1deg x 1deg grid. Estimating the spatial distribution of CO_2 sources is an important component of the modeling process used in direct measurements to derive estimates of emissions from measurements of CO_2 concentrations. This is discussed in detail in Section 5.

We note that the Vulcan Project has recently proposed to extend some of their methodology on a global scale to countries other than the U.S. They propose to use the assimilation approach described by Rayner et al. [15] called the Fossil Fuel Data Assimilation System (FFDAS). This system has the goal of assimilating many different types of data to produce spatially resolved global estimates of CO_2 emissions. Types of data could include: EIA global energy estimates, satellite observations of nightlights, Landscan population data, MODIS satellite data, ASTER thermal imaging data, and SCIAMACHY NO_2 data [12]. Aspects of this assimilation approach may be useful for the supply sector monitoring described earlier.

3.3 Estimating CO_2 Emissions using Knowledge of Energy Infrastructure

While the principal objective of monitoring energy infrastructure is to detect the signatures of a transition from a carbon-intensive to a low-carbon energy system, it will be useful to attempt to also use the monitoring information as "proxy data" to make an estimate of fossil fuel CO_2 emissions. Making

such proxy estimates has a number of challenges.

It is reasonably straightforward to design and set up a program to collect and evaluate proxy measurements, and we discussed some of the aspects in the preceding sections. In principle, at least for cooperative countries, most source locations of CO_2 emission are known and the technology employed is known, so in principle it is straightforward to design proxy measurements and calibrate them on known plants in the U.S. and elsewhere. In practice, proxy measurements are probably not an effective way to *survey* or *discover* sources of CO_2 emission, directly anthropogenic or indirectly anthropogenic, including those emissions from "Agriculture, Forestry and Other Land Uses" (AFOLU). Nor do proxy measurements of energy infrastructure provide much information about GHG other than CO_2.

There will surely be inaccuracies at the 20% level or greater from unreported or non-standard practices in less developed nations. Aspects such as IR flux that depends on emissivity, or wind turbines and solar farms that are placed at better-than-average, cherry-picked locations will not be easy to quantify accurately. It is also possible that large scale usage changes might occur, for example household, from rooftop photovoltaic power becoming dominant in India. This would cause a steady drift in accuracy of proxy estimates.

Despite these challenges, we recommend that estimates of fossil fuel CO_2 emissions using proxy data be attempted. As mentioned in the introduction, these proxy estimates will be useful for comparison with direct measurements and could also allow better process models to be built for CO_2 emissions that would be useful in the transport modeling required for direct measurement approaches.

4 DIRECT MEASUREMENTS

Our goals for this study are to prescribe an "investment market basket" of actions that would optimally measure CO_2 fluxes and whose accuracy we could evaluate. We believe there are a number of opportunities for better and more informative CO_2 measurements, and we will describe some potential capabilities, trying to provide estimates of what costs might be. However, we find that models are not yet mature enough to provide an estimate of the relative accuracy that different measurement strategies provide for flux estimation, and therefore we cannot yet optimize the value/cost proposition for different market baskets. For example,

- An OCO-3 [20] or "CarbonSat" [18] are being thought of as future satellites in LEO, but we think that placing a slightly modified system at GEO would be more valuable at the same cost.
- Relatively modest resources could be used quite differently to build an in-situ sampling network or fund aircraft sampling that would offer a remarkably dense set of measurements that we think might prove to be more valuable for inverse modeling from concentration to flux estimation.
- On the modeling front, CarbonTracker [7] is a well integrated system that derives ocean and biogenic CO_2 fluxes from inputs from meteorology, emission models, and sparse sampling of CO_2 concentration. We suspect, however, that CarbonTracker is close to being limited by its meteorological input and assumptions, and that substantial work in improved global modeling would reap large benefits in being able to more assimilate more information and more accurately quantify fluxes.

In this section we will describe various measurement opportunities in some detail, but we recommend that it is just as important to advance models that quantitatively assess the effectiveness of measurements as the measurement methods themselves.

The sections below discuss the context, methods, and constraints of CO_2 measurement, the opportunities for "*in-situ*" measurement, the role that satellite measurement can play, and finally recommendations. Appendices provide greater detail for the various sections.

4.1 Context

There are large, natural fluxes of CO_2 that have changed the planet's CO_2 level dramatically over history. Life has created large reservoirs of carbon captured into plants, soils, carbonates, and fossil carbon and oils, and the uptake and release of CO_2 continues at a scale that is large compared to anthropogenic emissions. Of course only a small change in greenhouse gas concentration may be needed to drive significant temperature and climate change, potentially causing a problem for an ecology that needs to support 10^{10} human beings. If we were only interested in the energy budget of the planet and how CO_2 concentrations might give rise to climate change, we might be contented with measurements from a single location, say Mauna Loa, because air mixes in longitude within a few weeks, across the equator in a year, and even through the stratosphere in less than a decade.

However, concern about the growing CO_2 concentration from anthropogenic sources spurs us to ask much more detailed questions about the location and fluxes from sources and sinks, both man made and natural. Even perfect knowledge of the CO_2 concentration everywhere does not tell us fluxes — additional information on diffusion rates and transport is essential. We therefore find ourselves needing to marry models of diffusion and transport, global meteorology, with concentration measurements.

Mixing a high concentration of CO_2 into ambient gas and transporting it creates entropy and involves non-linear, chaotic processes, so there are fundamental limits on how accurately it can be described. However, the growing field of "inverse modeling" tries to run the clock backward from sparse measurements of concentrations to plausible locations and fluxes of emission.

There are many heuristic formulae used for some of the chaotic and small scale physics that goes on, and as much as possible the models need to be constrained by data on pressure, temperature, wind velocity, humidity, and of course concentrations of the gasses of interest. The performance of these models is not bad and improving, but the greater the span in space and time they are asked to bridge, the less accurate they will be. Accordingly, the central question from the standpoint of this study, is:

What combination of spatial sampling density, temporal sampling density, location, and sensor performance parameters would best constrain estimates of the anthropogenic carbon emission from a region of interest, when the data are used in conjunction with an atmospheric transport model, meteorological data, and assumptions about sources and sinks?

4.1.1 Measurements

Any direct measurement samples the CO_2 concentration field, $C(x, y, z, t)$, over a pixel volume (dx, dy, dz, dt) at a particular location and time at (x_i, y_i, z_i, t_i). In-situ sensors sample a nearly infinitesimal volume, thereby being vulnerable to high spatial and temporal variability that may be impossible to capture in a model. Satellite sensors sample a volume with a ground sample distance and atmospheric column that dilutes CO_2 concentration by an order of magnitude relative to concentrations in the boundary layer and opens vulnerabilities to unmodeled effects such as clouds.

This challenge was studied by Olsen and Randerson (2004) [28], and Figure 10 shows their simulation of local and satellite data at a given location. The ground level plume from a large city or country might create a CO_2 enhancement of $\sim$5 ppm (parts per million mole fraction of dry air), and detection of that signal against natural variability using an in-situ sensor is not trivial, even given a 1 ppm accuracy. The background they calculate for a column CO_2 that a satellite would see is even more daunting, inasmuch as a sun-synchronous LEO orbit revisits a given ground location every $\sim$16 days,

and the column-integrated plume signal has only ∼0.5 ppm enhancement. Achieving a ∼1 ppm measurement from orbit has yet to be achieved.

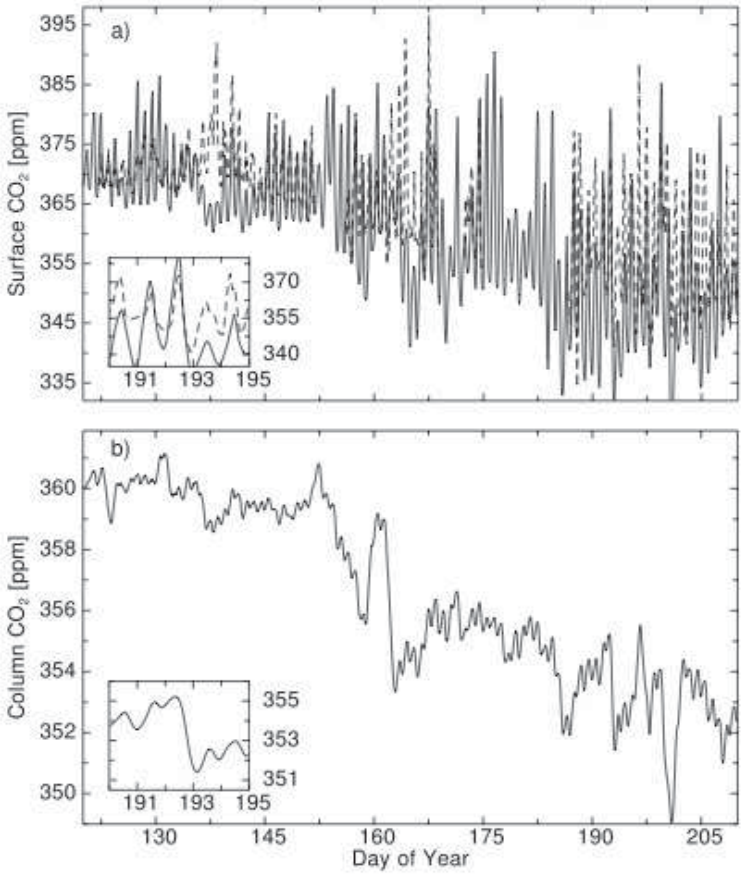


Figure 10: A comparison of simulated in-situ and satellite CO_2 data. The upper panel shows a simulation of expected signals from an in-situ sensor, at a height of 76 m, while the lower panel shows the expected signal as seen from orbit. Note the difference in scales. SOURCE: Olsen and Randerson (2004) [28]

Another example of the challenges faced by in-situ measurements from the RACOON network is shown in Figure 11 [30]. Similarly, we collected data in La Jolla during the study period (Appendix A) that reveal large CO_2 variations from local sources. Extracting anthropogenic signatures in the face of large background variations will be challenging for sparse density of concentration measurement.

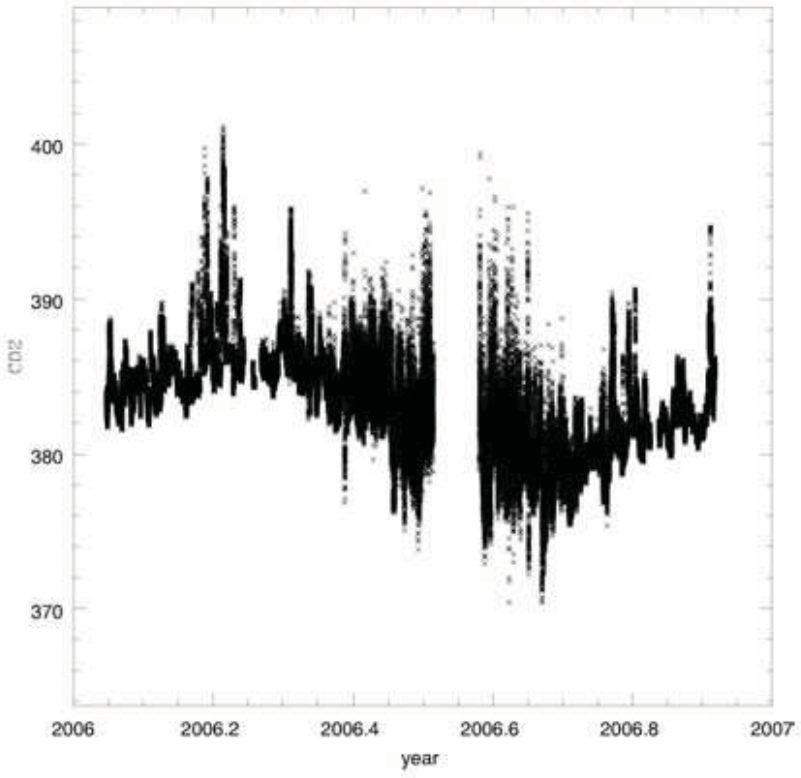


Figure 11: Variation in CO_2 concentration measured at the NWR station of the Rocky RACCOON network, for 2006. Note the wide variations seen on daily and seasonal time scales, which are highly significant compared to the measurement uncertainty of less than 1 ppm. The NWR site is on an especially barren ridge, in order to avoid diurnal CO_2 variations from vegetation as much as possible. SOURCE: `www.eol.ucar.edu/stephens/RACCOON/`.

Obviously a satellite's sampling trajectory of (x, y, z, t) is highly determined by its orbit. Atmospheric sensors as well can be deployed at a fixed site or on a mobile platform, and provide different trajectories in (x, y, z, t), with advantages and disadvantages to each. We can expect a spatial variability in CO_2 concentration similar to the temporal variability illustrated in Figures 10 and 11, and the (x, y, z, t) measurement trajectory entangles the spatial and temporal effects. A shrewd choice of trajectory may create information that is very informative to models, and a bad choice may be very ambiguous.

4.1.2 Coordination of measurements and models

As the importance of understanding CO_2 fluxes has grown, many groups have undertaken measurement and modeling efforts. To date, it is our impression that much of the work has been exploring the "state of the possible" without a great deal of coordination.

The web sites maintained by the CarbonTracker consortium [7] illustrate how inverse modelling is constrained by observations of CO_2. We believe that CarbonTracker is an excellent "state of the art production system" that provides a consistent, well defined product that has real utility. It also provides an excellent benchmark against which other production or R&D systems can be compared, and highlights the most important inverse modelling achievements and shortcomings.

The CarbonTracker modelling basically comprises a meteorological transport model (TM5) that integrates a nested global model based on the winds from the European Center for Medium range Weather Forecast (ECMWF), modules for ocean, fire, biosphere, and fossil fuel fluxes, concentration measurements from 89 sites, and a Bayesian assimilation that uses the likelihood of measurements for a given model and set of flux assumptions to adjust the flux rates.

Because of the global scale of their model and computational limitations, TM5 uses a global resolution of 6×4 degree (∼500 km) grid with nested regions at 3×2 degree, or 1×1 degree, along with 34 vertical levels. Their integration proceeds with an external time step of three hours.

The ocean module is based on 30 different ocean regions, each with a single parameter relating flux to Princeton/GFDL MOM3 ocean general circulation model. The fire module uses the Global Fire Emissions Database that is updated by information from sources such as the MODIS satellite observations of fires. The biosphere module uses the Carnegie-Ames Stanford Approach (CASA) model, and divides the world into 12 land regions, each comprised

of a mixture of 19 different ecosystems. Lastly the fossil fuel fluxes are taken from the EDGAR database. Only ocean and biosphere fluxes are fitted, so there are $30 + 19 \times 12 = 258$ adjustable parameters that CarbonTracker can use to bring the models into agreement with the observations. (In fact 113 are set to zero because of ecosystems not found in land regions or because some ecosystems have negligible CO_2 fluxes.)

CarbonTracker does a good job of parsing the global change of CO_2 into its distribution around the world, and provides interesting hints about how the oceans and biosphere are performing as sources and sinks. From the standpoint of anthropogenic emissions, however, CarbonTracker is essentially reflecting the EDGAR estimates. CarbonTracker is not designed to accept large and diverse CO_2 concentration measurements such as those from AIRS, it cannot simultaneously solve for meteorology and transport as well as CO_2 flux, nor does it have the spatial or temporal resolution to identify and quantify individual sources of CO_2 emission. We discuss more ambitious modeling efforts in Section 5.

For in-situ sampling, for example, actual measurement of CO_2 concentration was originally very expensive, making it cost effective to collect flask samples from about 50 sites that were brought to a central facility where the CO_2 could be measured by analytical chemistry or an accurate IR opacity instrument. This obviously limits the spatial and temporal density of samples, but was the foundation of the CarbonTracker inverse models. The network is now been extended to 89 locations, many performing the CO_2 measurement on site, using a gas chromatograph or an instrument such as a Picarro laser ring-down unit to calibrate against standardized cylinders of gas distributed through the World Meteorological Organization (WMO). These instruments do an excellent job of attaining <1ppm absolute accuracy, but cost over $100k and are non-portable especially with gas cylinders. CarbonTracker ingests only one measurement per site per day, and the analysis runs about two years behind the present.

A group at NCAR has developed an instrument called "AIRCOA" (Stephens et al. [29]) that uses a Licor LI-820 sensor (about $4K) and external calibration cylinders of CO_2 for a system cost of about $10k. They have deployed six towers in the Rocky Mountains separated by about 400 km (the "RACCOON" project) that sample CO_2 at $\sim$100 sec intervals, and attain a $\sim$1 ppm accuracy. This is easily accurate enough to map out the diurnal cycle of CO_2, and sampling at three elevations up their towers, they can see vertical CO_2 gradients. Only two of these sites are part of the CarbonTracker network, chosen to be particularly barren and therefore with low contamination from vegetation. The rapid time cadence, the contribution from vegetation, and the vertical gradient are not features that CarbonTracker can make use of at present.

It is our sense that capabilities have been advanced as *costs* have come down, but we do not yet know the *value* of the increased sampling. Does a relatively dense network of samples from the Rocky Mountains with a very dense temporal sampling improve our ability to ascertain fluxes from sources in California, for example? Does sampling the diurnal cycle help us understand the vegetation sink and source rates? Does it help calibrate the normalized differential vegetation indices (NDVI) estimates from satellites into CO_2 fluxes?

Another important question to which modelers need to provide input is the accuracy required for concentration measurements and the level of systematic drifts that can be tolerated. Although the locale of a source of emission may see tens of ppm enhancement in CO_2, as soon as the gas is mixed or when it is seen as part of an integrated column from a satellite or when it is part of a large ground pixel, the mean concentration can easily be only a few ppm or less. Therefore an accuracy of "1 ppm" absolute is commonly put forward as the threshold of where measurements become useful to models, with an understanding that much lower concentration uncertainty can be obtained by averaging.

We are concerned that satellite measurements may not achieve this. For example, Miller at al. [20] analyze the flux accuracy that ought to be achievable

by OCO and conclude that useful constraints should emerge if OCO achieves 1 ppm accuracy, but not at 5 ppm. However, the actual column-averaged CO_2 concentration downstream from even the largest cities and countries is a fraction of 1 ppm, indicating that the OCO analysis assumes considerable improvement from averaging before limitations from systematic error. Is this level of averaging strictly necessary? For example, changes in CO_2 concentration in space or time might provide excellent constraints for model's ability to obtain fluxes from concentrations, and satellites are likely to be able to achieve much lower systematic errors for such differences. Would this be informative to models or not?

Another possibility for management of satellite systematic error comes from the current best-fit decomposition of covariances in measurements (especially satellite) and reporting quantities such as "dry molar fraction of CO_2 integrated over an altitude kernel" and an uncertainty. It may be that models can help. For example, it may be that a satellite fundamentally measures a combination of CO_2 and water or pressure, with uncertainty as described by a covariance matrix. If the model has a better idea of water or pressure than the satellite observation, or derives constraints from a different mixture of the two quantities, the measurers are doing the modelers a disservice by not reporting the full covariance matrix and letting the modelers use it to evaluate their own likelihoods. The reason that this is so important is that satellite observations typically have very high signal to noise but are highly susceptible to covariance between molecular species, aerosols or other scatterers, ground albedos, contributions along the line of sight, etc.

We note that the comparison of model with data permits CarbonTracker to assess a bias and uncertainty distribution for each site, the errors and biases ranging from below 1 ppm to as much as $\sim$5 ppm. Does this mean that a site with a systematic offset of $\sim$5 ppm can be assimilated because of the consistency required by the model? Could a systematic drift of $\sim$5 ppm per year be tolerated? Clearly *some* sites must be known to be accurate, but perhaps a relatively small subset. Relief from 1 ppm absolute accuracy

would make an enormous difference for the cost, complexity, and weight of measurement equipment, and would open the door for much higher measurement density. Measurements must have some accuracy to be useful, but a better understanding of tolerable error margins is most important for measurement design.

What constitutes an optimum mix of measurement location and density? To what extent can models be formulated to remove systematic biases from measurements? These questions can only be answered by insight gained from models, and the optimum *value* per unit *cost* requires modelers and measurers to work together.

4.2 In-Situ Measurements

In this section we describe some of the sensor options for measuring CO_2 concentration within the atmosphere (or under sea), and follow with a number of deployment opportunities that may provide accuracy for inferred source and sink locations and fluxes.

4.2.1 Direct CO_2 measurement techniques

There are basically three methods for measuring the amount of CO_2 in a sample of gas: direct interactions with the molecules via analytic chemistry or gas chromatography, mass spectroscopy, and infrared light absorption or scattering.

van der Laan et al. [34] describe a modern automated system for high accuracy measurement of a variety of greenhouse gasses. The precision of the measurements is very high, $\sim$0.1 ppm, and the accuracy relative to standard cylinders from the World Meteorological Organization (WMO) can be better than $\sim$1 ppm. The automated system described by van der Laan et al. can make several measurements per hour. The equipment to make these measurements is relatively massive and requires clean lab space, but can be

obtained for a construction cost of ∼$200K and similar yearly operation cost.

This method has important but limited utility for very wide proliferation of sampling. The equipment is not portable, for example, and therefore requires flask samples for measuring CO_2 at altitude or at unimproved sites. As pointed out in Ref. ([35]), the accuracy of flask samples can be spoiled by momentary temporal variations in CO_2 concentration.

Mass spectroscopy systems range from accelerator mass spectrometers to conventional magnetic sector systems to small residual gas analyzers. An accelerator involves some work to introduce the sample into the vacuum of the spectrograph, but is capable of very high sensitivity for measuring rare species such as ^{14}C, for example. The instrument requires a large lab, and the costs are of order ∼$1M for construction and ∼$300K per year operation. We heard from Tom Guilderson [25] that the LLNL facility can measure 30 samples for ^{14}C per day, limited by counting statistics. For ^{14}C, at part per trillion relative to ^{12}C, there is no substitute. Conventional magnetic sensor systems measure the mass-to-charge ratio using an electromagnetic analyzer; it is typically used for isotopic analysis.

A residual gas analyzer that measures relative abundances of different molecules that are allowed to leak into the system can be relatively inexpensive (∼$10K) and compact. The precision of the measurement depends on effects such as diffusion rates, ionization susceptibilities, and molecular cracking, but it is plausible that concentrations of different isotopes of the same molecule, such as CO_2 or O_2, could be determined with quite high precision.

The third method exploits the characteristic infrared absorption spectra of molecules, using spectroscopy or non-dispersive IR (NDIR) measurements of this absorption to infer CO_2 concentration. Spectroscopy, either dispersive or Fourier transform spectroscopy (FTS) can be extremely powerful for disentangling different molecules, can offer very good quantitative accuracy, and is the normal choice for a satellite instrument. We will discuss its use at greater length in the satellite section, but for direct, *in-situ* measurements,

NDIR offers better performance per unit cost than spectroscopy. There are three tiers of NDIR instrumentation presently available, spanning a range in price and accuracy, and offering an interesting range of options that we describe below.

4.2.2 IR absorption in gas cells

Measuring CO_2 to a useful accuracy is doubly challenging because it is a minority constituent of air so the CO_2 signal is potentially small, (CO_2 concentration is only a few percent of typical atmospheric water concentration, for example, and water itself has only a few percent concentration overall), but also because we need to detect small relative changes in the signal itself (1 ppm out of a mean of ∼400 ppm). As a result it is extremely important to distinguish precision from accuracy; small systematic drifts or biases are important and are the limiting factor for instrumentation.

Non-dispersive infrared (NDIR) or cavity ring-down spectroscopy (CRDS) illuminate a sample of gas with an IR wavelength that is absorbed by CO_2, measures the attenuation, and relates that to the CO_2 concentration.

CO_2 has a number of vibrational bands corresponding to molecular stretches or bends, and these are split by many rotational lines. Two of the strongest bands are at 2349 $cm^{-1} \sim 4.26$ μm, corresponding to the asymmetric stretch mode (C moving opposite to the two O atoms), and at 667 $cm^{-1} \sim 15.0$ μm, corresponding to the bending mode. (The symmetric stretch mode with C stationary and the two O atoms moving in opposite directions has no dipole moment and is exceedingly weak.)

In a 300 K thermal background the asymmetric stretch mode is at a relatively dark wavelength, whereas the bending mode competes against ambient thermal radiation. The 4.26 μm band is therefore the normal choice for a gas cell. Other CO_2 bands include a strong overtone band at 2.75 μm that is overwhelmed by water absorption; a moderate strength overtone band at 2.0 μm that neighbors on the strong water line at 1.85 μm; and a relatively

weak overtone band at 1.6 μm that falls in a gap in water absorption. The 1.6 μm line is a good choice for satellite imagers that integrate a double pass through the entire atmosphere ($\sim$ 20 km path length). Figure 12 illustrates the HITRAN-2008 [32] line strengths of CO_2 and H_2O in the 1–20 μm range. Recalling the relative densities of H_2O and CO_2, the merits of the band at 4.26 μm become clear. The line strengths of the 4.26 μm band are illustrated in Figure 13.

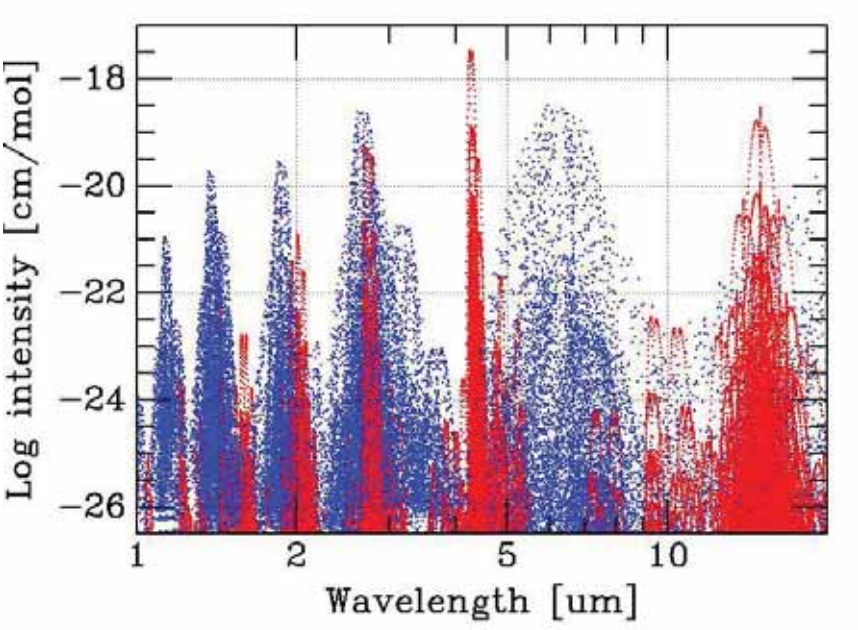


Figure 12: The CO_2 (red) and H_2O (blue) line intensities are shown between 1 and 20 μm. Each point is a single rotational line.

The absorption at a given wavelength by passage through a gas cell is given by

$$A(\lambda) = e^{-n\,L\,\sigma(\lambda)} \quad , \tag{4-2}$$

where n is the number density of absorbers, L is the effective path length through the cell, and $\sigma(\lambda)$ is the cross section for absorption.

The cross section is the product of the line intensity and the absorption profile. This profile is the relative probability that a photon of frequency ν in

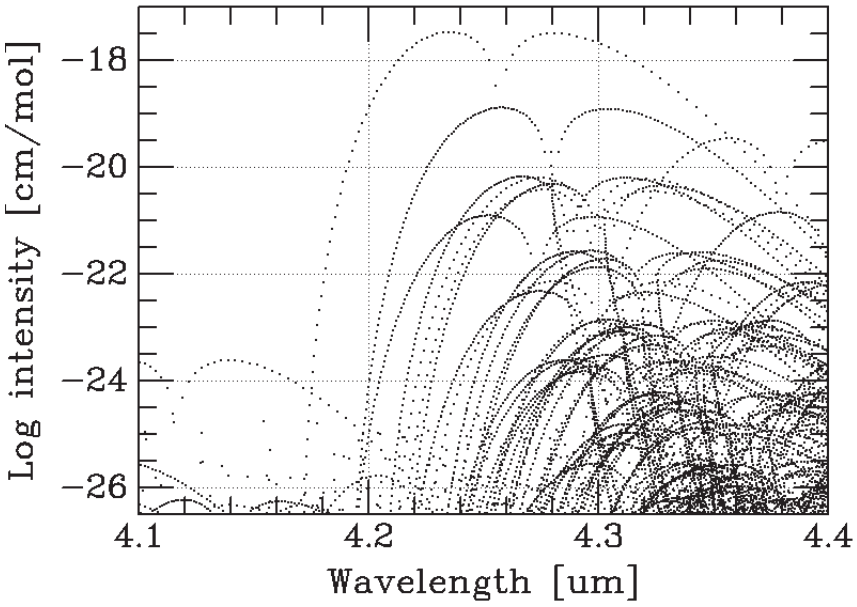


Figure 13: CO_2 lines and intensities are shown between 4.1 and 4.4 μm. Each dot corresponds to a single rotational line.

frequency interval $d\nu$ will be absorbed, and depends primarily on the Doppler shifts of the molecules and the state lifetime. At standard temperature and pressure (STP), CO_2 molecules have a typical velocity of 230 m/s, so the Doppler component of the line profile is essential a Gaussian of width $0.8 \times 10^{-6} \sim 0.002$ cm^{-1} at $\nu = 2500$ cm^{-1}. The typical molecular spacing at STP is $\sim$ 3 nm, the typical molecular size is $\sim$ 0.3 nm, so the mean molecular free path is $\sim$ 100 nm, and the mean time between collisions is $\sim$ 0.4 nsec, a frequency of $\gamma \sim 2.5$ GHz ~ 0.07 cm^{-1}. This establishes the lifetime of a given quantum state, and gives rise to a Lorentzian line profile component with this width, $P(\nu) \sim (\gamma^2 + (\nu - \nu_0)^2)^{-1}$. The convolution of these two functions is called a Voigt profile — a function that has a Gaussian core but extended $1/\nu^2$ wings. (The peculiar units of line intensity, cm/mol, finally become clear: it is cm$^{-1}\times$cm^2/molecule, absorption profile cross section per molecule integrated over frequency.)

If a line profile can be resolved, the total number density of absorbers can be retrieved by integrating the logarithm of the transmission over wavelength. If not, the dependence of net absorption in a finite bandpass on density of absorbers and path length is described by a "curve of growth". Generically this curve of growth for a single line depends linearly with nL when the exponential's argument is small, flattens dramatically into logarithmic increase when the Gaussian core of the line becomes saturated, and then resumes square root growth as the Lorentzian wings become significant. For CO_2 this behavior is complicated by the fact that there are many, closely spaced lines, and multiple families of lines at lower intensity, as illustrated in Figure 13.

Figure 14 illustrates the transmission through 1 m of air with 400 ppm of CO_2 at STP (based on notional profile parameters). Although the low resolution ($\sim$1.8 nm) spectrum has only 50% absorption, the individual lines are highly saturated and so the absorption will be a non-linear function of nL.

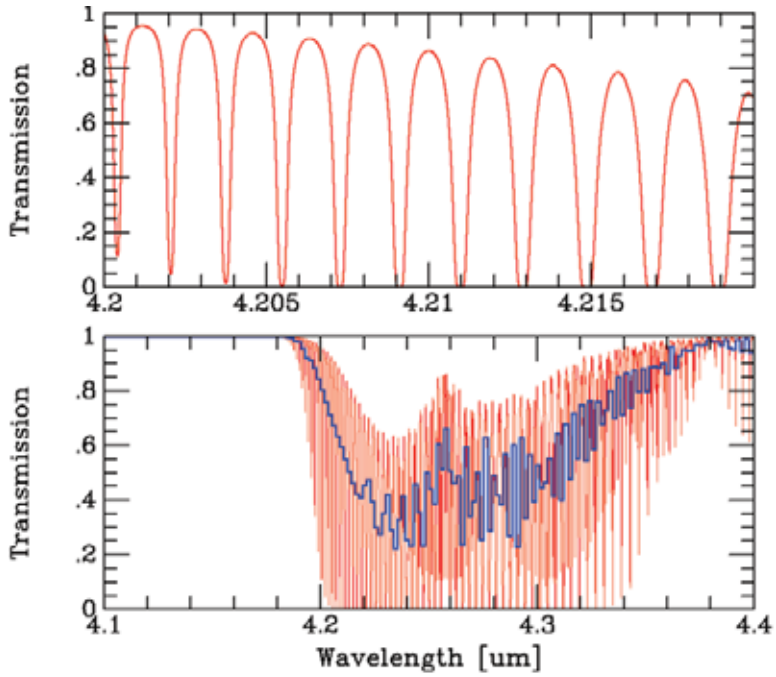


Figure 14: The transmission through 1 m of air at STP and 400 ppm of CO_2 is shown as a function of wavelength. The upper panel illustrates the saturation of the line cores. The lower panel shows the entire region and the mean transmission in 1.8 nm spectral bins.

A set of calculations with MODTRAN 5 illustrate in Figure 15 the curve of growth at STP for a square bandpass between 2300–2400 cm^{-1} as a function of CO_2 concentration, C, and path length, L. A similar curve could be produced for any bandpass or conditions desired.

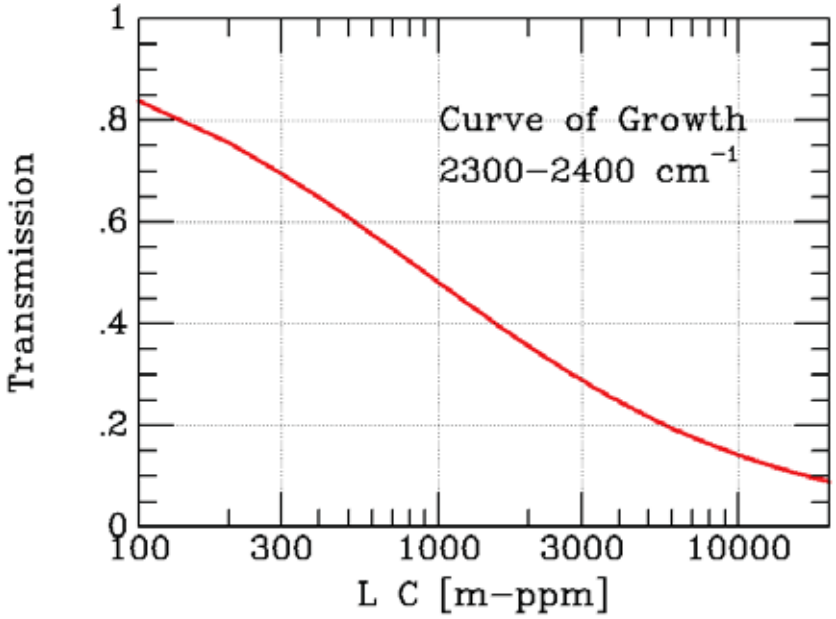


Figure 15: The transmission is a function of the nL product, and at STP this is a function of the product of CO_2 concentration, C, and path length, L. This curve is for a square bandpass spanning 2300–2400 cm^{-1} = 4.17–4.35 μm.

4.2.3 Cavity ring-down spectroscopy (CRDS)

Given a tunable laser that has a substantially narrower bandwidth than the $\sim$0.14 cm^{-1} full width of rotational lines, it is possible to scan across a CO_2 line and obtain the absorption profile. "Cavity ring-down" is an elegant technique that emits a short pulse of light into a high-finesse cavity, and monitors the decay in light intensity with time as it bounces back and forth. The ring-down time is relatively easy to measure and depends on the gas absorption and the cavity losses. By scanning on and off-line, the cavity

losses can be differentiated from the gas absorption, the spacing of lines being so close that the cavity losses are nearly constant. Since low absorption per pass is advantageous for timing accuracy and accuracy in obtaining density from the logarithm of the transmission, weak lines can be used as long as the cavity losses are low enough and the cavity is long enough.

Despite its elegance and ability to make an absolute measurement without reference to standards, CRDS places high demands on the quality and stability of the equipment. Examples of CRDS include the Picarro G1302 instrument, the Los Gatos Research "Fast Greenhouse Gas Analyzer", or the Tiger Optics "Prismatic". These units are typically about the size, weight, and ruggedness of a heavy suitcase, consume ∼100 W, cost about $50-100K, and provide continuous measurement with precision well below 1 ppm and accuracy of 1 ppm. Variations on these systems can also measure isotopic abundances (such as ^{13}C).

4.2.4 Non-dispersive infrared

It is possible to use a laser to scan individual rotational line absorption through a cell and measure intensity instead of ring-down time. For example, [33] describe a system based on a diode laser at 2.004 μm propagating through a beam splitter into a reference and Herriot cell (multi-pass cavity of length 26×8.8 cm), detected by InGaAs photodiodes. Intended for balloon measurements, their system achieves ∼1 ppm accuracy, uses a novel pressure matching scheme for their reference channel, weighs about 1 kg and consumes about 4 W. The instrument fits in a cylinder of diameter ∼20 cm and length ∼70 cm. Although a prototype, it could probably be reproduced for approximately $10K, much of the expense from the high precision laser.

If we give up on being able to resolve individual rotational lines, we can still measure the net transmission through a relatively broad bandpass (∼0.15 μm), centered on the 4.26 μm line. The basic measurement then involves a light source of some intensity I, passage through a gas cell of effective length L

(possibly involving many bounces), and detection of a signal S by a detector of gain G. Equation 4-2 provides a relation for the monochromatic transmission through a gas cell, and therefore relates S to the IG product. See also Eq. 4-3.

We use I and G in a generic sense because, although absolute calibration of I and G are possible, we would normally expect to include in the IG combination factors such as solid angles and reflectivities of the gas cell, and treat it as something to measure. As discussed above, if our system integrates transmission over a finite bandpass the exponential function must be replaced with a more complicated curve of growth g, and if individual lines are non-linear the dependence of cross section σ on temperature and pressure will affect the integral through the thermal broadening($\propto T^{1/2}$) and pressure broadening ($\propto nT^{1/2}$). This curve of growth and its dependence on P and T we would normally expect to calibrate by measuring system performance against gasses of known concentration as a function of P and T in domains of interest, rather than calculating ab initio.

The ideal gas law relates number density to pressure P, temperature T, and concentration C, $PC = nkT$, so we can write Equation 4-3 for the gas concentration,

$$S \approx IG\, g(\sigma' L C P/kT), \tag{4-3}$$

where σ' is a suitable average of σ over the bandpass. An example of g is shown in Figure 15.

There are therefore typically three unknowns that need to be measured in a linear system: the offset and gain of the S/IG combination, and the concentration C. (We assume P and T are known.) We would normally expect to obtain the offset by setting I to zero, the gain by arranging for the argument of the curve of growth and therefore the attenuation to be zero, and the concentration by a third measurement. There are a number of possibilities for causing the argument of the curve of growth to be zero without disturbing the IG combination:

a) change wavelength to where $\sigma(\lambda) \approx 0$,

b) reduce the CO_2 concentration $C \approx 0$, or

c) reduce the pressure P (in principle to zero, but reduction by a known amount allows extrapolation to zero).

The various NDIR instruments all use variants of these methods.

High quality NDIR instruments, typified by the Licor-7500 open path system or the Licor-840 cell, use a pair of temperature stabilized detectors that compare the transmission through a path in two $\sim$0.1 μm bands, one centered on 4.26 μm CO_2 absorption feature and the other at 3.95μm where there is minimal absorption. These systems are about the size and weight of a book, consume $\sim$5W in steady state, cost $\sim$\$5–10K, and provide continuous measurement with precision below 1 ppm and accuracy of 2 ppm. Many are in operation at relatively remote and unimproved sites, typically equipped with elaborate plumbing to convey air from multiple intakes or from calibrated gas cylinders. For example, Burns et al. [26] used a "man-portable" system whose net weight was $\sim$70 kg for their studies on calibration methodologies.

Inexpensive NDIR sensors, of which the SenseAir (www.senseair.se) or Gas Sensing Solutions (www.gassensing.co.uk) are examples, are available as OEM boards for $\sim$\$200 from distributors such as CO2Meter (co2meter.com). These systems use inexpensive IR emitters and detectors such as incandescent light and pyroelectric sensors equipped with narrow band filters, or IR LEDs and lead selenide, InGaAs, or InSb photodiodes. Although intended for applications such as HVAC where measurements to 10,000 ppm are needed, they do provide precision of 1 ppm with typical drifts of 10 ppm per day. These boards have a volume of about 10 cm^3, consume about 0.1 W, and can run unattended for years.

We evaluated three of these inexpensive units for several weeks, with results described in Appendix A. Although their shortcomings were readily apparent, it seemed to us that minor modifications could improve their accuracy ten-fold, and that they could be provided with an absolute calibration system that would be extremely light weight and low power. We therefore think that inexpensive NDIR could be brought to a state of maturity that would be valuable for detecting CO_2 concentration changes from sources and sinks, and could form a network providing a very high density of measurement.

4.2.5 ABET challenge

The Accreditation Board for Engineering and Technology (ABET) requires engineering students to design and execute a significant project that demonstrates their ability to meet real world needs. Many industries, foundations, and agencies offer funding to engineering design teams to come up with the best solution to a problem of real significance. We believe that the engineering of a CO_2 sensor that measures CO_2 concentration to 1 ppm, pressure to 100 Pa, temperature to 0.1 K, and humidity to 1% for a production cost of less than $500 is feasible, and in fact $200 should be achievable in quantity.

There are three distinct challenges:

- Design a clever labyrinth that achieves long path for small volume. (The SenseAir unit achieves about 10 cm in a 3 cm length; Appendix A describes how this could be increased by $\times 10$ for the same size.) This can be used in conjunction with existing, inexpensive NDIR boards to improve native precision to $\sim$1 ppm.
- Design a calibration unit that can deliver CO_2 at a known or calibratable concentration to 0.1% accuracy, weighs less than 30 g, and consumes less than 50 mW on average. Appendix A suggests a possible implementation.
- Build an integrated instrument that simultaneously delivers CO_2 concentration, P, T, and water vapor, q, to requisite accuracy, logs it with

an accurate time stamp, possibly a GPS location, with suitable communication: USB, cell phone, etc. This must meet the end goal of inexpensive mass production.

It strikes us that the pieces to successfully meet this challenge are all available, and the ingenuity, excitement, and competition of sharp college students would immediately result in a dramatically capable unit. The industry to mass produce and calibrate CO_2 sensors at the ~$100 level already exists, and it seems to us that the innovations that college students would come up with could rapidly be transitioned to production. If DOE truly were prepared to make use of thousands or tens of thousands of such units, we predict that they could be available in less than two years.

4.2.6 Sensor network

Given the existence of an inexpensive, accurate CO_2 sensor, there are many new opportunities for deployment that might be extremely informative for CO_2 fluxes. As always, we are speculating about the most valuable deployment and density, but we are confident that models can be developed that will point the way to the optimally informative array of sensors. Of course the existing, accurate, expensive sensor installations continue to be important, and can provide the absolute accuracy anchors for a model that assimilates all the data.

The most obvious deployment strategy is simply high, uniform density of CO_2 measurement near the surface, within the planetary boundary layer in which CO_2 concentrations are the highest. Whether this needs to be high above CO_2 emission and sinks, or whether it is possible to be immersed in their fluxes is an open question.

We consider cell phone towers to be an interesting place for CO_2 sensors. Obviously they can provide power and communications, and the density is very high because cell phone range is relatively short, especially in the higher

frequency bands. In the US there are some 140,000 cell phone towers, i.e. an average of one per 10 sq km. Although they in fact cluster with population density, that is also perhaps how CO_2 fluxes cluster as well. The rest of the world is tending to skip over the use of wires for telephone networks, and the predominance of cell phone communications is relatively higher than in the US. A CO_2 sensor that costs less than $500 could be a trivial part of the environmental suite of instruments that each cell phone tower already contains.

Another possibility for deployment would be ships or boats, for example sail boat masts or commercial shipping. We expect that virtually all vessels nowadays have GPS units, and most will have a GPS antenna at a reasonably high location. Installation of a small CO_2 sensor with the GPS antenna ought to be possible. A component of the ABET challenge would be a data logger, equipped with a few Gbyte of flash memory. At some 100 bytes per sample, and a sample every 10 seconds, the memory would provide storage for many years, easily adequate until an opportunity to connect to the internet arises and to dump to a central repository.

Experience working with the CIA's Crime and Narcotics Center has taught us of the difficulty in making 10% estimates of crop identification and vitality from remote sensing. Normalized difference vegetation indices (NDVI) and other multi- or hyper-spectral methods for vegetation estimation have proven utility, but the scale of vegetation CO_2 emission and absorption is so large compared to anthropogenic effects that it seems unlikely to us that remote sensing of NDVI alone will be able to constrain CO_2 fluxes at a useful accuracy. We wonder, however, whether an in-situ CO_2 monitor that can observe the diurnal variation in CO_2 might be able to calibrate indices such as NDVI. Spectral indices are mostly indicative of things like surface density of chlorophyll, lignin, underlying soil, and water concentration present in plant cells. The diurnal decrease in CO_2 concentration is directly related to the O_2 released in the process of fixing CO_2 from the air. A single point sensor cannot integrate the health and growth of vegetation over a large area, but it

could potentially calibrate areas that share a common set of spectral indices.

Given a sensor that costs only $200, say, it is worth considering deployments that end in the loss of the sensor. For example, a balloon, buoy, or floater that carries a CO_2 sensor equipped with an Iridium phone that provided a CO_2 measurement (along with time, position, temperature, humidity, pressure, etc) every 100 seconds for 3 hours could be considered a success by this metric. We can imagine, of course, a balloon or floater whose lifetime is measured in weeks, of course, and whose trajectory could cover hundreds or thousands of kilometers. Balloons can be designed to float at relatively low altitude, within the boundary layer, or could loft like weather balloons right into the stratosphere.

There are some 900 upper air stations that typically release two weather balloons per day. The radiosonde packages carried weigh about a quarter kilogram, and are in the process of being updated to include GPS receivers. The radiosondes are mostly lost, so of course they transmit their data back to the ground. The typical cost of an expendable radiosonde package is a few hundred dollars, so it is not inconceivable to include a CO_2 sensor as part of the package. The trajectory of weather balloons typically ascends at about 5 m/s to some 30 km before the balloon bursts, following a complex path according to the winds at different altitudes.

Wind turbines are becoming more and more common, both at a commercial size and scale as well as more modest units for individual residences. It seems that this might be another opportunity for siting inexpensive CO_2 sensors, since the turbines are purposely placed at an altitude that feels a wind with a large fetch, and they can provide power and communications.

Lastly, we do not discount the possibility of individual homeowners being willing to place a CO_2 sensor on a chimney or aerial and provide power and communications, simply to help understand the sources and sinks of CO_2 and their resulting effects on the climate. Reasonably inexpensive weather stations are also available for less than $200, for example the Oregon Scien-

tific WMR100, that can log wind speed, direction, temperature, humidity, pressure, and rainfall to a computer that can report results to the data repository. Obviously these data will have a wide variety in value, but we suspect that they could be informative to an appropriate model.

OCO will return some 150 million useful measurements of CO_2 over its lifetime, and at a cost of \$300M, such CO_2 measurements therefore cost \$2 apiece. (24 samples per second times a two year mission, diminished by a half for sunlight, and a fifth for land vs ocean and clouds.) For the cost similar to that of OCO, one could imagine deploying a network of a million sensors – one for each 20×20 km of the Earth's land surface. Would this be more or less valuable than the data from OCO? It requires appropriate modeling to be able to make the comparison, but we urge the modelers to "think big" and consider what could be learned from a truly dense array of sensors for CO_2, P, T, q, and possibly wind speed and direction.

4.2.7 Aerial

Appendix B describes possibilities of sampling CO_2 concentrations from aircraft. This could be an extension of the existing MOZAIC, CARIBIC, and IAGOS programs of measuring CO_2 and other gasses from commercial airliners, dedicated flights, or a new program similar to the "Open Skies" treaty that permits over flights to monitor nuclear weapons related activities.

The benefit of measuring CO_2 from aircraft is the ability to measure a vertical profile of CO_2, and therefore to provide constraints on the very difficult problem of meteorological convective transport. The shortcoming for commercial aircraft is that their trajectories sample from ground to stratosphere over a range of about 100 km from airports. It is not yet possible to place a value on what these concentration trajectories would confer to models.

It is striking, however, how cost effective programs involving commercial airlines can be. Let us assume that each aircraft charges a full-fare revenue of \$100k per year for carrying a 100 kg instrument package (approximately

the going rate for freight or passenger transport), and that we allocate $100k for a very capable instrument able to measure not only CO_2, but also H_2O, CO, CH_4, O_3, NO_x, NO_y, etc. For the cost similar to that of OCO we find that a fleet of 1000 aircraft could be equipped, and they cover a net ground track distance ten times as fast as OCO.

We do not mean by this to denigrate the value of OCO, but merely point out that the resources being dedicated to the measurement of greenhouse gasses are vast and can make massive improvements to existing programs. Exactly how the cost-benefit trade-offs work out need to be understood before committing resources, but may different opportunities exist.

4.2.8 Ocean

The ocean is a net sink of CO_2 and therefore is a crucial component in the balance of greenhouse gasses. Absorption of CO_2 is sensitive to ocean temperature and state, and the absorbed CO_2 leads to acidification, so there are climate feedback effects involving CO_2 and the ocean that are potentially important. Of course, the oceans border many countries and are therefore a useful vantage from which to identify and quantify fluxes of CO_2.

Appendix C elaborates on the opportunities we see in equipping ocean floats and gliders with CO_2 sensors. The Argo program is already under way to equip the oceans with $\sim$3000 floats, and is an opportunity to provide a dense inventory of CO_2 levels in the ocean, if even a fraction of the floats were equipped to measure CO_2 and carbon particulates.

The appendix also considers the possibilities of measuring CO_2 from ships plying shipping lanes, and discusses the delicate issues arising when such measurements can be used for treaty verification. On the one hand, ships could automatically register atmospheric CO_2 concentration as part of normal, meteorological measurements, as described above in the section on sensor networks, and the results could provide input to a comprehensive model that quantifies CO_2 sources and sinks as well as meteorology and transport.

On the other hand, such results might not be accepted as evidence of treaty non-compliance, and indeed, the situation is unclear as to whether ocean and atmospheric sampling is allowed outside of territorial waters but inside Exclusive Economic Zones.

4.3 Isotopes

Appendix D describes possibilities we envision that exploit isotopes of carbon and oxygen. The most straightforward is measurement of ^{14}C, the radioactive isotope that is in all living organisms but which has decayed from fossil fuels. ^{14}C exists only at the part per trillion level, and requires relatively expensive accelerator mass spectroscopy to measure. Since each ^{14}C measurement costs about $100 it cannot replace more mundane sampling of ^{12}C for quantifying sources and sinks of CO_2, but shrewd choices for ^{14}C measurement can disambiguate fossil fuel emission from other CO_2 fluxes.

A different opportunity may arise by measurement of the relative abundances of the stable oxygen and carbon isotopes: ^{16}O, ^{17}O, ^{18}O, ^{12}C, and ^{13}C. These can potentially be assayed with a relatively inexpensive residual gas analyzer, and they offer some clues to the vegetation and combustion origins of CO_2. The ratio of $^{47}CO_2$ to $^{44}CO_2$, for example, is nominally 44 ppm, but is affected by biological pathways and combustion temperatures. (See [37] for a recent discussion of isotopic tracer methods, as well as use of CO as a tracer).

Ideally one might expect that measurements of CO_2 and models could point to a particular location as a source of CO_2 and quantify its flux. Further discrimination of CO_2 into natural versus fossil fuel fractions could then be assayed with appropriate samples depending on whether models or treaty requirements deem it important or not. ^{14}C or other stable isotopic measurements are likely to be more useful for detailed diagnosis than for broad search and mapping of CO_2.

Compared with direct measurements of CO_2 concentration, many of these isotopic tracers have the advantage that plants[1] rather than flasks or artificial sensors can be used as samplers. Plant growth naturally averages isotopic ratios over an entire growth season, thereby suppressing many short-term fluctuations, so that a single plant sample may replace dozens of flask or sensor measurements for the purpose of reducing statistical noise. As noted in Section 5.2, some modeling approaches (needed to infer fluxes from concentrations) average transport processes over atmospheric timescales of order weeks and hence are naturally compatible with measurements averaged over similar timescales. Some signatures, such as the $^{14}C/^{12}C$ ratio, are almost uniquely associated with fossil-fuel burning, whereas CO_2 concentrations vary for many other reasons, as discussed elsewhere in this report. The prices paid for these advantages of isotopic measurements are (i) additional processing steps required to separate the isotopes from one another and from other components of the plant matter; and (ii) the need to understand and correct for biases in the signatures due to diurnal, meteorological, and seasonal modulations in plant respiration and photosynthesis.

4.4 Satellite Measurements of CO_2

4.4.1 Introduction

An instrument in orbit can monitor CO_2 over wide regions of land and ocean, larger than would be practical using in-situ measurements, including entire countries which may or may not be cooperative. Monitoring CO_2 from space involves measuring either the attenuation or thermal emission of infrared light at the specific wavelengths where CO_2 molecules have a large cross section. This is done either by observing sunlight reflected from the Earth's surface, or by using the thermal emission from the Earth's atmosphere. Several CO_2 measuring instruments are already in orbit and another, the Orbiting Carbon Observatory (OCO-2, Kuang et al. [19] and Miller et al. [20]) is scheduled

[1] including, perhaps, agricultural exports in the case of non-cooperative countries, provided that the places where the crops were grown can be adequately determined

for launch in 2013, as summarized in Table 3. These are all in low Earth orbits (LEO), and one of the questions we investigate in this section is the utility of adding one or more CO_2 instruments in a geosynchronous Earth orbit (GEO) or higher.

Table 3: Existing and Imminent (OCO-2 Scheduled 2013) Space-based CO_2 Instruments. All are in Low Earth Orbits (LEO).

Method	Instrument	CO_2 Measurement	Precision	GSD
Reflected Sunlight	SCIAMACHY	Total Column	3-10 ppm	30 km
Reflected Sunlight	GOSAT	Total Column	4 ppm	10 km
Reflected Sunlight	OCO	Total Column	1 ppm	1.5 km
Thermal Emission	AIRS	Mid-troposphere	1-2 ppm	13 km
Thermal Emission	IASI	Mid-troposphere	38 ppm	$\sim$100 km
Thermal Emission	TES	Mid-troposphere	$\sim$5 ppm	$\sim$50 km

There are several challenges that need to be overcome to obtain a useful measure of CO_2 concentration, including:

- Absorption or thermal emission by molecules other than CO_2 (especially H_2O, which is much more abundant) must be disentangled.
- The return seen in a given pixel depends on ground albedo and may be contaminated by other, variable sources such as clouds, aerosols, etc.
- The signal depends on state variables such as pressure and temperature.
- The ground sample distance (GSD) must be small enough to meet the desired goals, e.g. $\sim$2 km to measure localized sources/sinks such as power plant emissions, etc.
- The instrumental stability and calibration chain must support the requisite accuracy and precision.

We discuss these challenges in detail in subsections below. As we will see two key issues will be a) the high spectral resolution needed to permit good discrimination of CO_2 from other molecules, aerosols, and ground albedo,

and b) the small GSD needed to permit better spatial resolution of CO_2 processes and more frequent seeing through openings between clouds.

In order to achieve good wavelength discrimination for item a), orbiting instruments employ one of two methods: dispersing the light over a number of pixels with a grating spectrometer as in OCO, or instead measuring the wavelength with an interferometer using Fourier Transform Spectroscopy (FTS) as in GOSAT. The size and cost of a high resolution, stable spectrograph or interferometer does not present the impediment for a satellite that it would for in-situ sampling. On the other hand, the direct illumination across a defined volume of gas that is key for in-situ instrumentation is extremely difficult for satellites. (The proposed ASCENDS mission proposal seeks to do this with LIDAR in the 2020 timeframe.)

For both items a) and b), higher spectral and spatial resolution requires more collecting area and more pixels or more time to achieve the requisite signal to noise ratio (SNR) over a given total scan area.

A very important aspect of satellite measurements is that they average CO_2 over a significant volume of atmosphere (their "integration kernel"). This averaging implies that satellites see CO_2 concentration changes from localized sources that are an order of magnitude smaller than an in-situ sample within the boundary layer, and the satellite measurements must therefore be more accurate. Another challenge is that the trajectory of a LEO satellite, at least, provides only an instantaneous snapshot of CO_2 concentration along the ground track. The inability to average in time implies a susceptibility to unmodeled diurnal or small scale variations in CO_2, as illustrated in Figure 10.

Table 4 lists some of the orbit alternatives available for CO_2 monitoring from space. This is not meant as a complete list, but rather is provided to illustrate the various trade-offs one can make.

Table 4: Orbit Choices for CO_2 Flux Determination

Orbit	Pros	Cons
LEO ∼500 km 90 min period	good ground resolution instruments in orbit already measurements at nadir	mixes spatial and temporal var. variable scattering geometry
Sun-sync circular 500–1000 km ∼ 100 min period	polar orbit, global coverage limited range of scatt. geometries measurements at nadir	mixes temporal and spatial var. revisit ∼16 days (for 10 km swath)
Geosynchronous 36,000 km 24 hr period	near-synchronous observations can monitor entire diurnal cycle single satellite sees China & India	limited/no access to high latitudes needs larger aperture, vs. LEO limited longitudinal coverage
"Molniya" 40,000 km apogee 12 hr period	long dwell at high latitude some diurnal sampling high latitude access	further away than GEO inclination poor for equator limited longitudinal coverage
"Tundra" 46,000 km apogee 24 hr period	long dwell at high latitude fixed scattering geometry good diurnal sampling	limited area coverage further away than GEO inclination poor for equator
L1 1,500,000 km Stationkeeping	perpetual dwell over illuminated area fixed illumination geometry whole-Earth, diurnal sampling	limited access to high latitudes much poorer ground resolution and much lower bandwidth than GEO

4.4.2 Measuring atmospheric CO_2 by reflected sunlight

All present and planned satellite measurements of CO_2 absorption utilize the same basic measurement strategy exploiting the 3 different wavelength bands shown in Figure 16. The weak CO_2 overtone band at 1.6 μm, which has sizable (30%) column absorption and nearly linear sensitivity to CO_2 column density, serves as the main CO_2 sensor. The weak O_2 M1 absorption band at 0.76 μm (the A-band) calibrates air pressure and temperature. The stronger overtone CO_2 band at 2.06 μm calibrates the water vapor background and, together with the 0.76 μm band, addresses wavelength-dependent variations of surface albedo and airborne cloud and aerosol particles. Further discrimination against the latter effects is provided if the instrument wavelength resolution is sufficient to resolve the individual components of the bands, particularly the 1.6 μm band and thereby "chop out" the more uniform background.

A key aspect of the transitions at all three wavelengths is their inherently good line spacing that permits individual line components to be seen. Each transition has the same simple band structure; two rotational branches with resolved lines formed by $\Delta J = +1$ and -1 transitions between ground and excited state rotational levels, with rotational spacing of about 1 cm^{-1} (30GHz), and widths of individual lines in the lower atmosphere determined by pressure broadening of about 5 GHz. The atmospheric temperature determines the energy of the rotational levels occupied and hence the number of rotational lines that appear in the spectrum. (The change in spacing of these lines seen in the figure is due to the difference in rotational constant in the ground and excited states.) With such resolution, as noted above, combining the 3 lines enables the state variables such as temperature to be determined, as well as a host of modeling parameters such as the extra broadening due to water vapor, the difference in pressure broadening vs. the change in atmospheric pressure with height, the effects of line saturation, and the albedo and aerosol contributions between lines.

Satellites are typically designed to obtain adequate SNR from the reflection off of ground. The reflection signal from water at SWIR wavelengths is only 2%, a factor of 10 lower than sand/soil and a factor of 3–5 lower than vegetation. Over ocean the plan for OCO is to use sun-glint mode to get

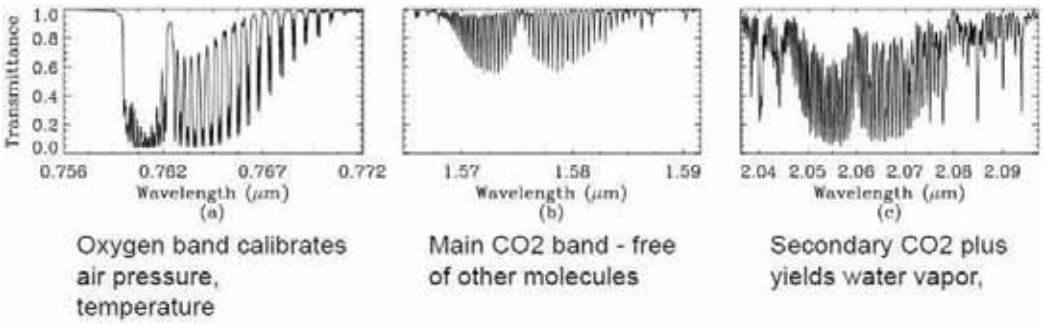


Figure 16: This figure plots the transmission through one atmosphere, in the spectral regions used for monitoring CO_2 and other gases from space, using reflected sunlight. SOURCE: [22]

adequate signal. This strategy is shown in Figure 17. OCO will be able to switch between nadir and sun-glint modes, and after initial testing in space may be programed in favor of one or the other as the dominant mode.

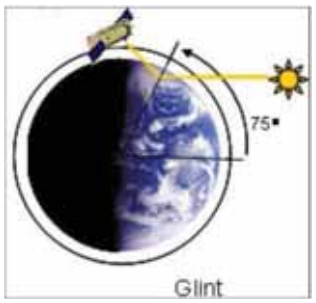


Figure 17: This figure shows the LEO satellite CO_2 monitoring coverage for nadir viewing over land and sun-glint viewing over ocean.

The retrievals outlined here, either over land or ocean, require validation for each satellite. For this purpose a validation network, called the Total Carbon Column Observing Network, TCCON, is in place, operational, and has been calibrated to the WMO Standard [38]. Target observations over TCCON sites will be a critical component of a validation strategy for future satellite instruments.

4.4.3 Potential for monitoring power plant emission

In a treaty environment, an important goal for future generation satellites is to constrain anthropogenic CO_2 emissions by detecting and quantifying strong, localized CO_2 point sources such as power plants (PPs). Although OCO will have the requisite GSD and CO_2 sensitivity, it will not normally have sufficient swath width to capture the usual factory plume. In order to achieve this, the results of an analysis by Bohvensmann et al. [18] indicate that the single ground pixel CO_2 column retrieval precision needs to be better than 1% as this is the expected order of the CO_2 column enhancement relative to the background. A number of other criteria need to be fulfilled, in particular the satellite's swath width needs to be sufficiently large to achieve

frequent mapping of PPs and their surroundings, and the GSD must be small enough (<2 km) not to dilute the PP plume and to have a good probability of a cloud-free scene. (Miller et al. [20] estimate this probability from MODIS observations to be approximately 35% $(\text{GSD}/1\text{km})^{-0.5}$.)

Bohvensmann et al. (2010) put forward a promising satellite concept called CarbonSat that has slightly less spectral dispersion (0.11 nm vs 0.08 nm), GSD (2 km vs 1.7 km), and SNR than OCO, but which has a 500 km swath width (compared to 10 km for OCO). Shown in Figure 18 is a simulated retrieval of a PP plume by CarbonSat. An emission of $F = 13.00$ MtCO_2/yr was used to generate the data seen by CarbonSat. The inversion result was $F = 13.22$ MtCO_2/yr.

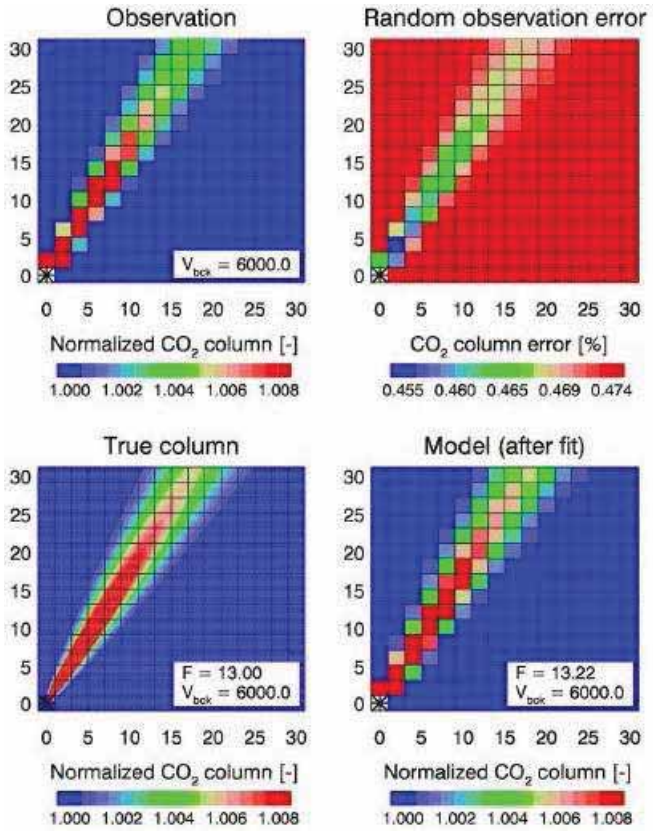


Figure 18: This figure shows the retrieval of a power plant plume by carbonsat as described in the text (from Bohvensmann et al. [18])

CarbonSat is proposed as a LEO satellite, but it is important to note that only 0.5 cm aperture is needed to capture enough photons from the 2 km GSD for the required SNR. Exactly the same spectrograph could be flown at GEO with 25 cm aperture fore-optics with exactly the same capability, except that at GEO the satellite could scan or dwell where ever is most informative, over the $\sim$ 10 hour span of daylight. The CarbonSat proposal is somewhat notional and there are other factors besides raw photon collection that drive instrument design, but any real CarbonSat design could also be deployed at GEO with fore-optics that are $\times 40$ larger. Inasmuch as the spectrographs of SCIAMACHY and OCO are about 0.5 m^3 and 150 kg, it seems that CarbonSat would be about the same and therefore considerably larger and more massive than any required fore-optics. We discuss the relative merits of GEO versus LEO at greater length below.

4.4.4 Thermal IR measurements of atmospheric CO_2

An alternative way to measure atmospheric CO_2 from space is by its thermal emission. The most precise results have been obtained using the Atmospheric InfraRed Sounder (AIRS, concentration retrievals discussed by Chahine et al. [21]) launched in 2002. AIRS is a cross-track scanning grating spectrometer with 2378 channels and spectral resolving power of about 1200 extending from 3.7 to 15.4 μm with a 13.5 km field of view at nadir. The AIRS spectrum is shown in Figure 19. AIRS utilizes a companion microwave sounder, AMSU to retrieve temperature, etc. in the presence of clouds on a horizontal scale of one AMSU field of view or 45×45 km at nadir, the equivalent of 3×3 AIRS footprints.

The directly measured quantity in AIRS is the upwelling radiance at the detector for all the channels in the spectrum. AIRS retrieves variables of interest X_i (for example the mixing ratios of atmospheric gases) by the method of "vanishing partial derivatives". In this method, for independent variables X_i, the values selected are those which make the partial derivatives $\partial G(n)/\partial X_i$ vanish individually. Here the residual $G(n)$ at the nth iteration is defined as

the sum of squares of the difference between the measured radiances and the radiances derived through forward calculations using the retrieved geophysical state of the atmosphere. Therefore, even though the observed spectra may not differentiate between the contributions of individual lines, the partial derivatives can.

For a given gas, a judicious choice of wavelength range is made to obtain the optimum line-strength, because the interaction probability along the column determines the "averaging kernel" for thermal emission measurements. For a strong enough line that the radiation is absorbed in much less than the height of the atmosphere, the averaging kernal is located near the top of the atmosphere, whereas for much weaker absorption the averaging kernal moves down towards the earth surface, In the case of CO_2, weighting functions in the troposphere are shown in Figure 20 for three wavelength ranges:

700–720 cm^{-1} with moderate absorption emphasizing the mid troposphere.

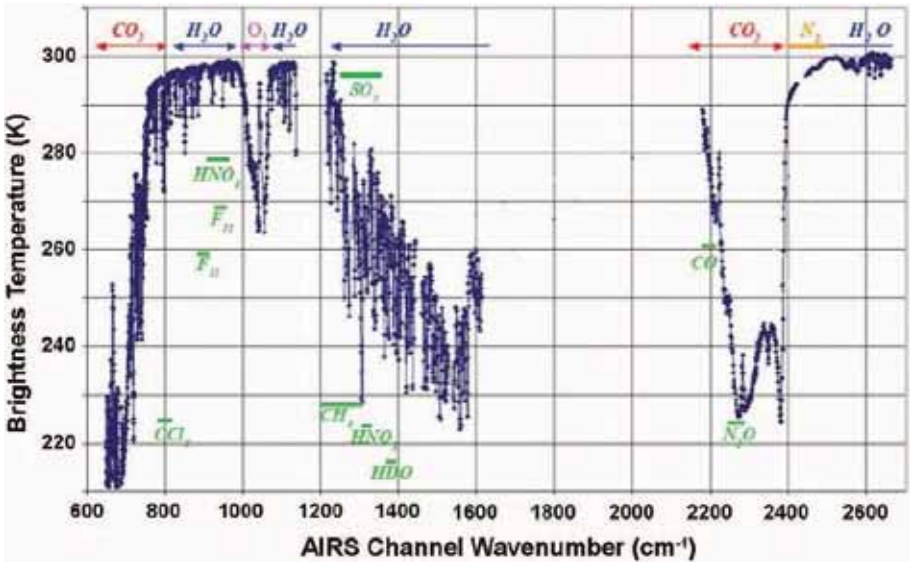


Figure 19: This figure plots the transmission through one atmosphere, in the spectral region used by AIRS to observe thermal emission from the atmosphere and the Earth's surface.

These measurements have been carried out already.

$\leq$700 cm^{-1} with strong absorption emphasizing the upper troposphere. These measurements are planned for 2010.

740–790 cm^{-1} with weak absorption emphasizing the lower troposphere. These measurements are planned for 2011.

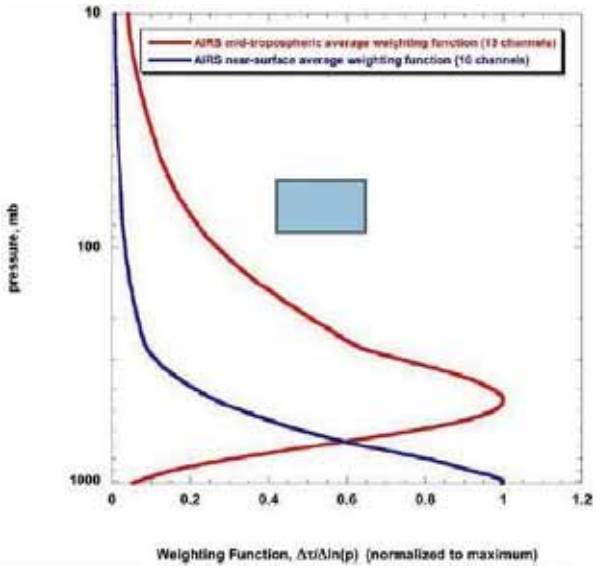


Figure 20: This figure plots the weighting curves that can be used by AIRS to observe CO2 concentration at various heights in the troposphere.

It is noteworthy that the mid-tropospheric CO_2 results agree with aircraft data in regions where the measurements overlap (Chahine et al. [21], Matsueda et al. [23]), and AIRS obtains the same 5 ppm seasonal variation to 1.2 ppm. The thermal IR thus offers an important means of obtaining global CO_2 concentrations in the mid troposphere (and likely in the upper troposphere too), that can be usefully compared with GOSAT and OCO-2 measurements of total column CO_2. Thermal IR also offers the important advantage of providing measurements both day and night. It remains to be seen with what precision AIRS will also be able to measure the lower tropospheric CO_2. It is thought that it will be possible to obtain the temperature

accuracy needed to reject the surface to the level of 2 ppm CO_2, but rejecting water vapor may prove more difficult.

Thermal IR measurements of CO_2 should be considered as an important option for a GEO orbit which would offer similar advantages in coverage as in the case of reflected sunlight absorption measurements already discussed.

4.4.5 The Merits of geosynchronous orbit

We find the idea of monitoring CO_2 from geosynchronous orbit to have numerous advantages. One of the most important is that this approach largely decouples the temporal from the spatial structure functions of CO_2 variability. This ability to take a series of instantaneous "snapshots" of CO_2 across an entire country provides a time series of maps of columnar concentration. We strongly suspect that this data set will be less sensitive to systematic errors in modeling and the inversion from concentrations to fluxes, than the spare spatio-temporal sampling from LEO instruments. It is important to remember that the repeat interval at which nadir-looking LEO instruments return to the same patch of ground is of order 15-20 days, depending on orbit details (even assuming cooperative clouds).

While the decoupling of temporal and spatial structure is the principal advantage of a GEO satellite there are several additional important advantages:

- GEO offers the option, unavailable at LEO, of revisiting important locations frequently on time scales of hours or days, and looking at them longer. As already noted, this option would be key to monitoring power plants and other sites with anthropogenic sources of CO_2.
- Since a number of LEO instruments exist already, with plans to be augmented soon by OCO-2, the GEO instrument would provide a nicely complementary spatio-temporal sampling function. For example, the high latitude regions where a GEO system observes at high slant receive dense coverage from LEO systems in nearly polar orbits.

- Launches of communication satellites into GEO orbits happen at the rate of about one per month, so we suspect that a "piggyback" instrument could possibly be a cost-effective approach, sharing the power and communications infrastructure from the comsat bus.
- It goes without saying that simultaneous data are extremely powerful for improving systematic errors. Equipping a GEO satellite with both a see-to-ground CO_2 sounder at 1.6μm as well as an IR imager akin to AIRS that collects CO_2 measurements from IR emissivities is a combination that should be carefully considered. Measurements of GHG other than CO_2 would also be advantageous.

There is no technical reason that a satellite cannot perform as well or better at GEO than at LEO. For example, consider what would happen if OCO -3 were flown at GEO (OCO-GEO) without any modification whatsoever (other than pointing mechanism, bus, communications, etc., of course). OCO-2's detector pixels map to a 64 m GSD on the ground at LEO, which would become 3.2 km pixels at GEO. These pixels are normally binned by 20 in the spatial direction at LEO, but could be returned unbinned from GEO, for a net, instantaneous FOV of 3×500 km. The OCO optics are not optimized for imagery in the spatial direction, so the GSD would be blurred to $\sim$10 km in the spatial direction, but we comment on that issue below.

Since OCO-GEO would be looking at the same surface brightness Earth as OCO-2, it would achieve the same SNR per pixel in the same exposure time of 0.33 sec. However, the pixels would not be binned by $\times$20 in the spatial direction, so the exposure time (or number of added exposures) would have to be increased by $\times$20 to achieve the same SNR per GSD. OCO-GEO would therefore observe some 250 km^2 per second, as compared to OCO-2's 70 km^2 per second. With a 10 km^2 pixel the probability of OCO-GEO obtaining a random cloud-free sounding is 18% instead of 25% for OCO-2 at LEO. Of course OCO-GEO can direct its observation to locations that meteorological satellites verify are cloud-free, so in principle OCO-GEO can have a very high duty cycle for successful soundings. OCO-GEO pixels over ocean can

be binned by 3×3 to obtain adequate SNR with a 10 km GSD.

Continuing this example, there are a number of obvious improvements that could be made to "OCO-GEO" to make it an even more capable instrument at GEO.

- Reoptimize OCO's optics to perform better in the spatial direction so that the spatial pixels are unblurred.
- OCO uses the H1RG devices that are $\times 10$ lower noise than the nominal detectors assumed by CarbonSat, but only uses $\sim$190 pixels in the spatial direction. Fill the entire 1024 pixels for a swath width of 3000 km.
- OCO's telescope aperture is 11 cm. Double it to 22 cm for a 1.6 km pixel size and a swath width of 1500 km.
- (More challenging) Increase the aperture to 25 cm and decrease the f-ratio of the spectrograph camera to f/1.6 for an exposure time reduction of 30%.
- (More challenging) Use H2RG detectors for OCO-GEO and fill the entire detector for a swath width of 3000 km with 1.6 km pixels. The additional pixels in the dispersion direction could be used for CH_4 at 1.66 μm or perhaps imaging both polarizations separately.

The first three (minor) modifications would lead to a collection rate of 350 km^2/sec to 1500 km^2/sec for 1.6 km or 3.2 km GSD, i.e. the area of the United States in 1.5–6 hours. The latter two modifications would provide 1.6 km mapping of the US in about 1 hour. There are many other optimizations such as slit width and spectral resolution that can decrease exposure times further.

This example is by no means the best that can be achieved with a GEO satellite that uses a spectrograph, and by this example we are not literally advocating that the OCO-III instrument be flown at GEO. It does illustrate, however, that measuring CO_2 from GEO is technically no more difficult than from LEO. In many ways GEO is a simpler location to design for because the portion of Earth at reasonable slant angle only subtends about 10 degrees,

and that is comparable to the efficient range of blaze angle from a grating. Therefore there is a wide design space for a system with wide swath and efficient spectrograph. In addition, the radiance from Earth is high enough that a 25 cm aperture at GEO is adequate to obtain the required SNR at high dispersion, ∼2 km GSD, and ∼1 sec exposures. Since the high spectral resolution requires a ∼10 cm beam onto a grating anyway, there is no advantage in being closer to Earth — this is why OCO's aperture is 11 cm and why OCO has such high magnification and therefore bins by ×20.

A grating is not the only way to obtain a spectrum. Duren (2010)[24] described the PanFTS instrument for the GEO-CAPE mission proposed to NASA. This ambitious instrument covers 0.26 μm to 15 μm with 7 km nadir GSD and 0.05 cm^{-1} (0.03 nm at 1.6 μm) spectral resolution. (A narrow field channel has 0.25 km GSD but far too little spectral resolution to measure CO_2.) It can collect a 900×900 km^2 image every 67 seconds, for a coverage rate of 12000 km^2/sec (600 km^2/sec if fore-optics were changed to provide 1.6 km pixels).

The objectives of such a mission would go beyond what we consider in this report and would include measuring species such as pollutants (O_3, NO_2, NH_3, SO_2, CO, etc), and greenhouse gases (CO_2 CH_4, N_2O, CO, etc). The notion of doing precision trace gas monitoring from GEO has been discussed by Burrows et al. [27] and Orphal et al. [31]. The GEO-CAPE and PanFTS mission, as proposed, is "something for everyone", does not have the technical readiness level of OCO, and is not particularly optimized for CO_2 or other GHG measurement. However, it would be straightforward to optimize for GHG measurement in support of treaty verification, and the simultaneous use of SWIR for detection of CO_2 absorption and thermal IR for CO_2 thermal emission is a potent combination.

We believe that, on balance, the properties of GEO enable powerful opportunities for classes of CO_2 sounder design that are not available at LEO. We summarize the advantages and disadvantages of GEO vs. LEO. Advantages include:

- A piggy-back ride on communications satellite lowers cost, and provides essentially infinite bandwidth and power.
- Ground sample size is largely dictated by the number of detector pixels flown and the desired revisit cadence.
- Long exposure times or binning is possible over ocean.
- Synoptic coverage possible, enabling detection of diurnal CO_2 variations, peeking through clouds, monitoring weather changes that move CO_2 around, monitoring CO_2 sources and sinks on a timescale that is meaningful.
- Geosynchronous orbits are possible that drift up and down in latitude (e.g. "tundra orbits"), with the potential of providing alternate view angles at the satellite's longitude, although it may make the communications aspect more problematic.

Disadvantages include:

- A GEO satellite views ground pixels off-nadir. We note however that a) LEO satellite systems successfully obtain CO_2 concentrations in the face of off-nadir solar illumination, b) pixel size is not much smaller than the atmospheric scale height so the integration volume is not greatly larger than the GSD, and c) a wandering geosynchronous orbit could alternate nadir views at each latitude.
- A GEO imager only sees one third of the world at once, and three would be required for global coverage.
- High latitudes are basically invisible (although we note that that is where LEO satellite coverage becomes very dense).

We believe that a GEO sounder offers an extremely important complement to existing and planned LEO satellites. Modelers need to weigh in on the relative value of the temporal coverage and other GEO advantages, and make recommendations. In parallel, satellite designers should carry out detailed

calculations of the accuracy expected from off-nadir imaging and changing sun angle, and need to think how to optimize the design of a GEO imager.

4.4.6 L1 monitoring of CO_2

It is also worth considering placing a satellite at the Lagrange point, L1, about 1.5 million kilometers towards the Sun from the Earth. Being 40 times more distant than GEO, it is not practical to maintain the same ground sample distance with the same f-ratio, so some combination of increase of ground sample distance and f-ratio would have to be implemented.

If we imagine for sake of argument that our "OCO-GEO" example were equipped with an 0.8 m telescope and placed at L1, it would have a 20 km ground sample distance (and full-planet swath) but otherwise be the same in terms of exposure time and CO_2 measurement accuracy. Available downlink bandwidth would be significantly less than at GEO, which might place additional constraints on operations.

However, this satellite would stare continuously at the sunlit face of the Earth, and it would not have the foreshortening of a GEO vantage, so all points on the Earth at latitudes lower than ∼50 degrees would be visible for approximately 6 hours, from 9 AM to 3 PM every day. The cloud contamination within 20×20 km is potentially a serious issue (Miller et al. [20] assess the probability of cloud-free at 12%), but a scan of the entire sunlit face of the planet would take less than an hour, so a given spot has 6 chances per day of being seen without cloud. As with a GEO satellite, CO_2 over dark ocean can be easily measured by simply binning 3×3 soundings. Of course we think of an L1 satellite as a complement to pre-existing LEO (and GEO) satellites, but we want to emphasize that relatively low spatial resolution required for CO_2 measurement can be obtained from many different locations.

The "Deep Space Climate Observatory" (Triana) is a $100M satellite proposed in 1998 to view Earth from L1 with an imager and radiometer for climate monitoring purposes. The imager incorporates a 30 cm telescope

and 10 channels from 317 nm to 905 nm to obtain 8 km nadir GSD over the face of the planet every hour. It has been on hold because of NASA prioritization and its perceived scientific value (c.f. [39]), but it serves to illustrate that L1 is an attainable location and that many of the practical difficulties can be addressed.

4.5 Findings and Recommendations: Direct Measurements

In this section we have reviewed the state of CO_2 measurement, both in-situ and satellite. We have tried to elucidate costs and values for different methods and deployments, but we find that we are not able to carry out a real recommendation of an "optimal market basket" because models do not yet exist to quantify flux uncertainties for a given set of measurements. We are optimistic that such models can be brought to some maturity in a relatively short period of time, which we discuss in the next section.

Nevertheless, there are two opportunities that to us appear to have a certain pay-off. The first is development of a low cost, 1 ppm accuracy sensor and calibration system that could be very widely deployed and networked. The second is CO_2 measurement from geosynchronous orbit or beyond. Neither of these is necessarily superior to the existing, sparse, in-situ networks or satellites in LEO, but as a complement and next step, we find them compelling.

To illustrate the trade-offs we face in establishing an optimal network of in-situ sensors, we observe that the following options would each be available for an investment of around $20M (7% of the cost of the OCO reflight):

1. Acquiring and deploying about 100 high-accuracy cavity ring-down CO_2 sensors. This would roughly double the global sampling network of the existing Carbon Tracker sites.
2. Equipping 100 commercial or military aircraft with high quality instru-

mentation and paying one year, full freight fare. A single airliner flies about a million miles per year, so this is some 160 million km of ground track measurement per year, almost exactly the distance that OCO will cover per year.

3. Equipping 500 ARGO floats with CO_2 and C particulate sensors for directly evaluating ocean sinks and source of CO_2.
4. Acquiring 1,000 stabilized NDIR sensors akin to AIRCOA, with <1 ppm accuracy. This would amount to about one hundredfold expansion of the Rocky RACCOON network.
5. Designing, acquiring, and installing a network of 20,000 inexpensive, few-ppm accuracy sensors. Options include cell phone towers, (possibly sacrificial) balloon flights, ships, or other favorable locations.

Our suspicion is that a much higher density of sampling may not improve our understanding of the overall global burden of CO_2 and climate change, but will be extremely informative for identifying and quantifying the sources and sinks of CO_2.

We advocate investment in developing a low cost, high accuracy CO_2 sensor right away (with auxiliary measurements, attendant communications, and assimilation). Investment in understanding the various deployment opportunities and costs right now is worthwhile — it is possible to ascertain the *costs* and what information will be delivered with some accuracy right now.

For satellites, we find that there are no fundamental technical roadblocks to deployment at GEO or beyond. Indeed, GEO may offer some synergies with the Com-sat industry that provide financial leverage for greenhouse gas monitoring capabilities. The temporal and spatial coverage afforded from a GEO (or beyond) vantage point seem to be a most valuable complement to LEO instruments such as SCIAMACHY, AIRS, and OCO.

We therefore recommend immediate design studies of new or adapted satellite concepts for GEO or other high orbits, and exploration of Com-sat partner-

ships, with a goal of launch in 2018 or earlier. The outcome again should be a clear identification of the capabilities and costs of different configurations.

In a few years time, with appropriate investment in models it should be possible to understand the *value* of all the various opportunities and then to make an investment that is optimal for the purposes of determining greenhouse case fluxes and treaty verification.

5 MODELING/ASSIMILATION/INVERSION

5.1 The Flux Inversion Problem

5.1.1 Introduction

Whether carried out by in situ sensors or by satellite, measurements of greenhouse gasses yield values of the concentration averaged over some instrument response profile. Profiles can be local in space (point sensors) or spatial-averaging (remote sensing pixels), local in time or temporally averaging. The first task, then, of any emission-monitoring program is to infer fluxes of gases from these concentration measurements. We have already discussed some of the issues involved in the use of models to infer fluxes from concentration measurements in Section 4.1.2.

In this section, we focus on the state of the art in carrying out the aforementioned flux inversion. Before proceeding, it is worth contrasting two different uses of these algorithms. In a tracer release scenario, there are assumed to be no natural sources of the material in question; often we can also assume we have good estimates of the sinks, either occurring in the bulk (e.g. natural decay of radioactive tracers) or by deposition onto the ground. In contrast, many of the important greenhouse gases have large natural sources and sinks. This means that calculating the fluxes is only the first step in an overall effort to determine changes in anthropogenic sources. In particular, one will have to develop a good understanding of natural variability so as to be able to place bounds on specifically *anthropogenic* fluxes as opposed to total fluxes.

The simplest formulation of the flux inversion problem is arrived at by assuming that all transport processes are linear in the relevant concentration variables. Under most cases, this seems to be a reasonable assumption as the most critical processes such as advection by winds and mixing in the vertical direction are indeed linear. Then, a vector of measurement μ_i, $i = 1, N_m$ is

typically taken to be

$$\mu_i = \sum_a H_{ia}\sigma_a \ + \ \epsilon_i \tag{5-4}$$

where σ_i is a set of N_f fluxes (which in principle should also include initial data at the beginning f the monitored period), ϵ_i are measurement errors and H is a $N_m \times N_f$ matrix which summarizes how sources are connected to measurements via transport equations. An example of the output from this type of calculation is shown in Fig 21. The indices respectively indicate how many measurements take place and run over both spatial locations and time values. Source indices represent unknowns and again indices encode both spatial location and time period. To give one concrete example, in a model with simple diffusion, an averaged concentration over a region of volume V centered at $\vec{x}$ and temporal duration T centered at time t, would be given for a single source as

$$\int_V d\vec{x'} \int_T dt' \int d\vec{y} \int d\tau \ G(\vec{x} + \vec{x'} - \vec{y}; t + t' - \tau) \ \sigma(\vec{y}, \tau) \tag{5-5}$$

where G is the usual Greens function for an advection-diffusion equation

$$\frac{\partial G}{\partial t} - D\vec{\nabla}^2 G + v(\vec{x}, t) \cdot \vec{\nabla} G \ = \ \delta(\vec{x} - \vec{x}')\delta(t - t') \tag{5-6}$$

Here $\vec{v}(x, t)$ is the advecting velocity.

Before writing down the solution of the inversion problem, it is worth mentioning the issue of spatiotemporal discretization. In order to solve the transport question, space and time must be discretized. The space and time scales chosen for this discretization need have no direct connection with the space and time scales chosen for the measurements and/or the fluxes being solved for, other than the obvious requirement that one clearly cannot hope to resolve fluxes on scales smaller than the solution grid. We will return later to the issue of how various choices of discretization can affect the inversion outcomes.

The standard Baysian approach to finding fluxes relies on minimizing the error by proper choice of the σ_a subject to a priori information. A standard assumption is that the ϵ_i are normally distributed with covariance $\mathbf{R}$. If

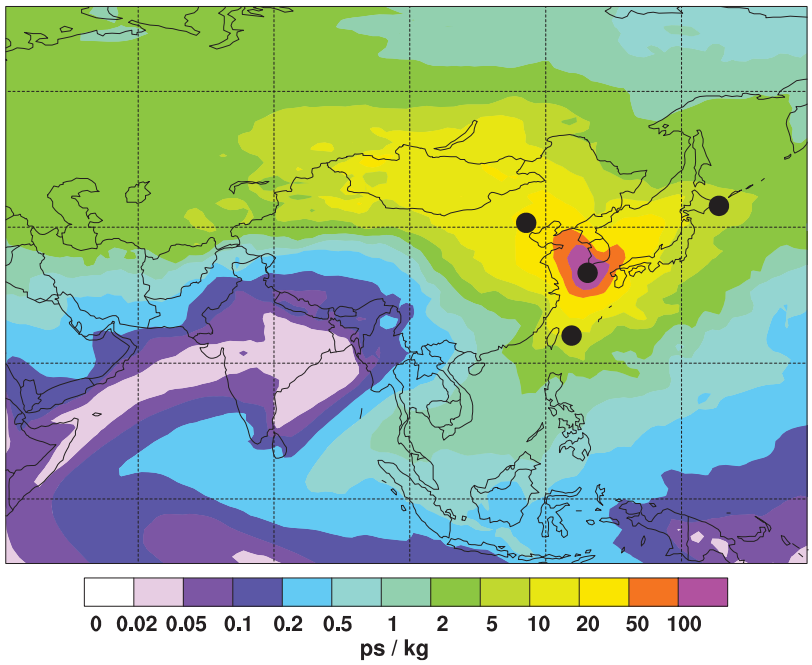


Figure 21: Footprint emission sensitivity in picoseconds per kilogram obtained from backward calculations using an advection-diffusion model averaged over all model cal for all stations and for the entire year 2008. Measurement sites are marked with black dots. from [40]

one also assumes that the prior distribution of the sources is Gaussian with covariance $\mathbf{B}$, we must minimize

$$(\mu - \mathbf{H}\sigma)^T \mathbf{R^{-1}} (\mu - \mathbf{H}\sigma) + \sigma^{\mathbf{T}} \mathbf{B^{-1}} \sigma \tag{5-7}$$

This assumes that the prior distribution has zero mean, but this condition can easily be removed. The solution of the problem is

$$\sigma_{est} = \mathbf{B}\mathbf{H}^{\mathbf{T}} \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}}\right)^{-1} \mu \tag{5-8}$$

and the variance of the estimate is found to be

$$V_{est} = \left(\mathbf{B^{-1}} + \mathbf{H^T}\mathbf{R^{-1}}\mathbf{H}\right)^{-1} \tag{5-9}$$

has clearly been reduced from $\mathbf{B}$ by means of the measurements.

One can extend this framework to account for prior distributions that are not Gaussian. This is definitely important for the tracer release problem since the sources must be positive and occur at only a few isolated sites. Bocquet and co-workers have shown how one can extend this Bayesian framework to allow for such non-Gaussian processes using a maximum entropy principle [44]. Details are available in a series of publications. Others seem to use ad hoc procedures such as manually adjusting fluxes to zero if the algorithm gives negative values [40], or parameterizing the fluxes in such a way as they are intrinsic positive (taking the flux to be $\exp\sigma$, for example [45]). This latter approach has the distinct disadvantage of making the problem non-linear in the flux variable and making error estimates much more difficult.

An extensive discussion of the formalism for data assimilation and modeling inversion is given in Appendix E.

5.2 Discretization Issues

The aforementioned framework is used in the greenhouse gas monitoring community in a variety of manners. In some works, data is taken from existing ground monitoring networks or from existing satellites and inverted to determine actual fluxes. In a second protocol, forward models with assumed sources are used to create artificial data for proposed systems and used to evaluate their added utility in reducing uncertainty. It is critical to realize that neither of these two uses typically investigates modeling errors, for example in the vertical transport physics, perhaps due to insufficiently accurate weather information. We will return to this issue later and assume for the moment that the model, as written down as a partial differential eqation, is accurate.

An early example of this type of study is that of Rayner and OBrien [41]. Here, the goal was to assess the utility of a nominal remote sensing as compared to the global-view monitoring network. The authors assume that the flux can be discretized into a total of 26 regions (covering the entire globe); a

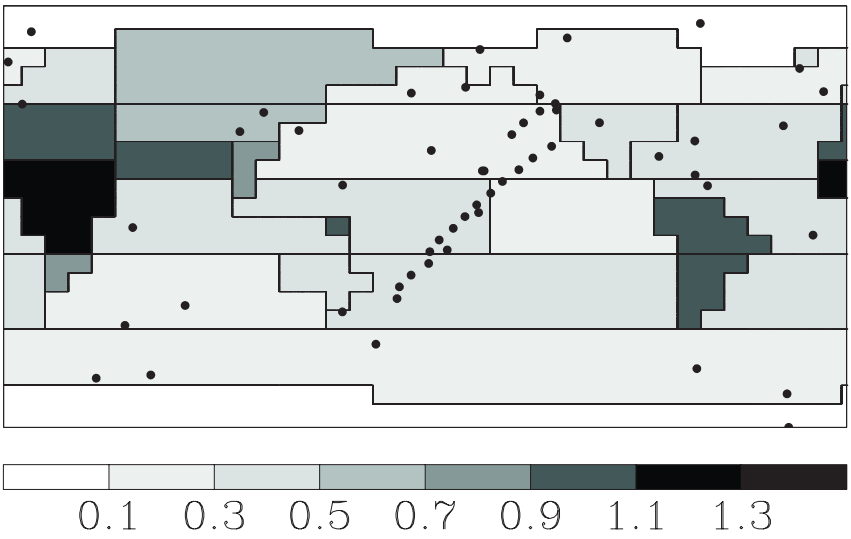


Figure 22: The 26 region flux inversion grid (polygons) and the 56 mainly surface stations. Stations are marked by dots. The shading represents the a priori standard deviation of sources (GtC yr $^{-1}$). From [41].

picture of their discretization is shown in Fig. 22. In the time domain, their movements are averaged over one month and similarly the fluxes are taken to be constant over each month.

In this paper, the authors are concerned with the needed accuracy of satellite measurements so as to reduce flux inversion uncertainty. To this end, they assign initial flux errors which are chosen to be of order 1.2 GtC/yr. (over land regions) and similar values over different oceans. They first show, using the aforementioned formulas, that the ground-based measurements appeared to reduce the uncertainties to approximately .5 GtC/yr. over North America, slightly worse over Europe. These numbers need to be compared to the total fossil fuel based emissions from Europe estimated to be around 1 GtC/yr. Clearly, this resolution, even if taken at face value, means that the ground system could not detect a change in emissions of less than 50-100%. This falls far short of what might be desired.

It is crucial to recognize that the limitation revealed by this analysis is not due to what normally would be called measurement error. The data uncertainty in the ground sensors in the neighborhood of 1 ppmv are, as discussed in this paper, much larger than the precision of the instruments. Instead, the estimated error comes from the extreme variability of the actual measurement data which then needs to be averaged over one month to get even close to something that can be meaningfully compared to model predictions. The failure to reproduce short time scale variability with a coarse-grained model is hardly surprising. First, we have assumed constant fluxes over very large regions and over long periods of time. This is mismatched to data taken at isolated points in space and with a short averaging period. An explicit example of this mismatch can be seen in the work by Stohl [40] on inverting for the sources of hydrochlorofluorocarbons (HCFCs) in east Asia; the best fit to emissions on a 1^0 x 1^0 degree cannot reproduce the spike in the actual data (see Fig 23). This is attributed both to the coarseness of the emission grid and the exactly matched coarseness in the model, which also assumes a 1^0 x $1\ ^0$ grid.

In general, then, we have several choices that must be made. At the most basic, we have already mentioned that the transport equations must be discretized. In practice, discretization is determined by available computational power and available meteorological data. There is no guarantee that for any given problem, these artificial constraints will coincide with the actual desired outcome, namely that the solution of the discretized problem be close enough to the behavior of the actual physical system being modeled (recall that we are assuming that we have an accurate continuum model). This is a well-known issue in fields such as computational fluid mechanics, where subgrid models have been developed to try to ameliorate the problems engendered by solving the relevant flow equations on coarse grids that cannot properly resolve the physics.

To explain the problem, let us imagine we only care about emissions over country-scale regions. One might theoretically imagine that solving the trans-

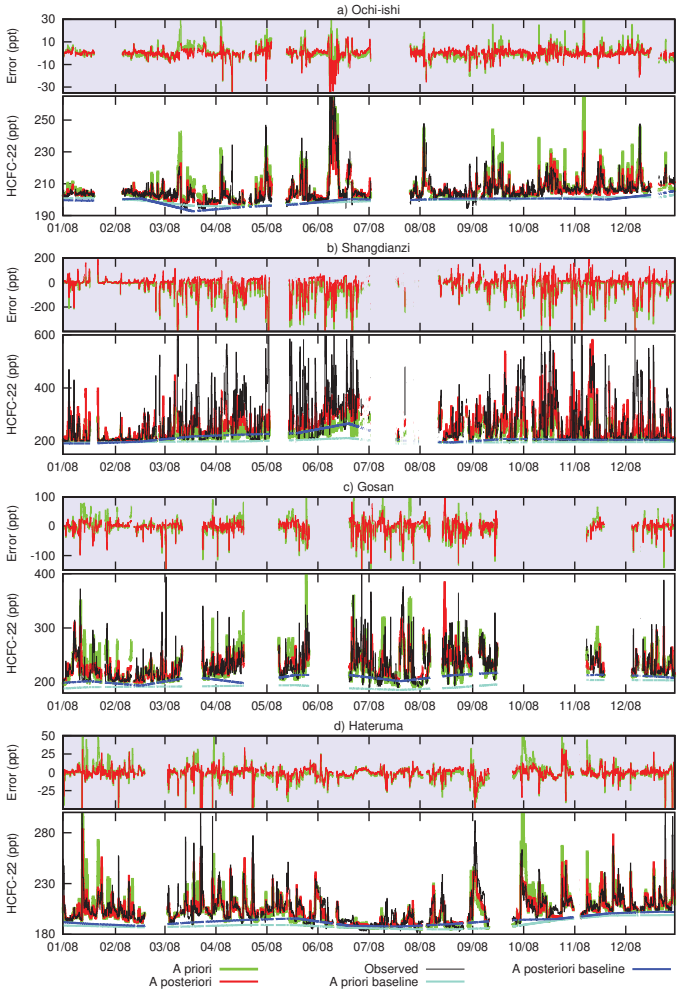


Figure 23: HCFC-22 time series for monitoring stations at (a) Ochi-ishi, (b) Shangdianzi, (c) Gosan, and (d) Hateruma. For every station, the lower panels show the observed (black lines) as well as the modeled mixing ratios using the a priori emissions (green lines) and the a posteriori emissions (red lines), the a priori baseline (cyan lines), and the a posteriori baseline (blue lines). The upper panels show the model errors based on the a priori emissions (green lines) and the a posteriori emissions (red lines). The spikes at Shangdianzi cannot be captured as the measurements are too close to large sources in Beijing and hence are subject to large fluctuations caused by small-scale transport events. From [40].

port equation on say, a 3^0 x 3^0 grid would be good enough to detect relevant flux sources. But, there is no guarantee that the transport from one coarse cell to a neighboring coarse cell is predictable from just 3^0 x $3^{\,0}$ degree information about concentration, wind, etc. Subgrid correlations between wind velocity and concentration could easily lead to cell to cell transport even if the mean value of concentrations in the two cells on the coarse scale are equal. Such effects should appear, at the very least, as stochastic terms in the model, but the real issue is the extent to which the structure of the fluctuation depends on the fields themselves (multiplicative as opposed to additive noise) and can lead to systematic problems. It is clear, also, that similar remarks apply in the time domain, where substep correlations can lead to biases in the transport predicted by the coarse-grained model.

The second level of discretization involves the size of the inferred fluxes. In general, it seems like a dangerous notion to take this scale to be the same as that of the model discretization; this does not allow any determination of whether or not the space-time characteristics of the data can actually be explained by slowly-varying fluxes. Results for fluxes would be more convincing if one had the freedom to vary the flux discretization within a *fixed* model discretization and then discover the relevant scale of the source distribution, rather than impose it from the start.

In practice, the amount and type of measurement data places constraints on the flux inversion. First it is reasonable that one will not get a meaningful inversion without having, say, an order of magnitude more data points than unknown flux variables. In the work of Stohl [40], 11706 observations were used to determine 1377 unknowns. But, this is not just a numbers game. Certain types of fluxes will not leave any significant trace on certain measurements, and conversely if there are some variables where net effect is not seen in any significant subset of the data, the inversion will be ill-conditioned. For other fluxes, the algorithm should just return a fitted value determined by the a priori input information. Thus, producing an emission sensitive plot

of the type already shown in Figure 21 is an essential part of the process and should have been calculated for any proposed system.

The counting of measurements is thus somewhat tricky. Measurements taken over very short intervals are not really independently useful if the model uses meteorological data averaged over much longer time scales. Measurements separated by short distances will not give independent information about fluxes being determined at a much coarser level of discretization.

This short excursion into the many inter-woven issues involved in the flux inversion problem should indicate that the situation can be quite complex. This is perhaps an area that could be sorted out with the help of focused computational studies in which large- scale resources could be used to create a fine-grained simulation which could then be investigated by a series of coarse-grained reductions. This type of study would decouple the problem of model accuracy with that of computational and measurement resolution. As far as we can tell, the community has not yet undertaken such studies.

5.3 Satellite Measurements

As already mentioned, one of the uses of inversion has been to understand how proposed systems could be used to reduce emissions uncertainty. In the paper already mentioned by Rayner and OBrien [41], they concluded that a nominal satellite system could be competitive with ground level observations if measurement errors would be less than a few ppmv. A more detailed study focusing on the specific OCO measurement footprint [43] led to the conclusion that the error would indeed be reduced to .1 to .2 Gtc/yr. over regions the size of Europe. If this is true, they would indeed to good enough to monitor a 20% change in atmospheric emissions, assuming of course a constant natural background. They predicted significant reduction in the error, by about a factor of 2, over India and China (see Figure 24). But, this study was carried out on a rather coarse grid (3.75^0 x 2.5^0) and no attention was paid to the extent that models at this scale may not be all that predictive. In fact,

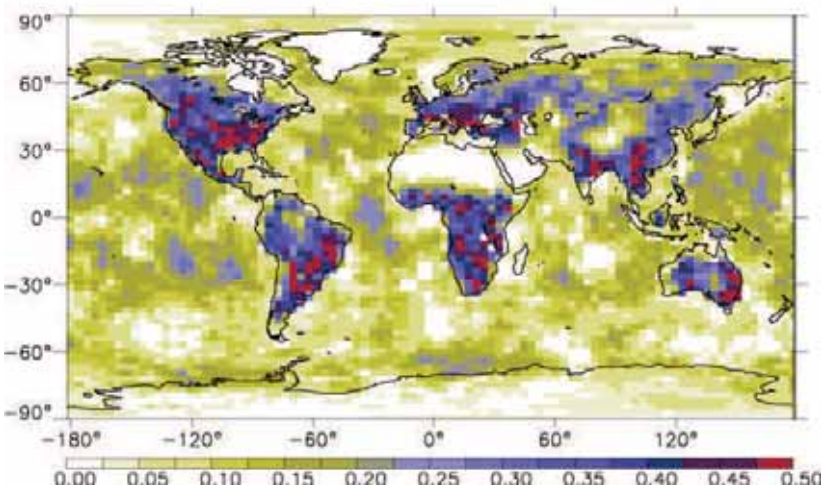


Figure 24: Fractional error reduction of the monthly mean grid point CO2 surface fluxes. The error reduction is defined as (1- σ_a/σ_b), where σ_a is the posterior error standard deviation and σ_b is the prior error standard deviation. From [43].

it would have been impossible to do so, as they took their meteorological data from GCM models rather than from weather service observation. This limited the resolution to GCM scales.

One message that is unmistakable in the paper by Chevallier et al [43] and the work of Rayner and others [41] is the reliance on a priori estimates. If the goal of a carbon monitoring system is to ensure 20% measurement accuracy without any additional input, none of these papers offers much encouragement. If the goal is to test deviations from inventory reports that can be mostly trusted, and if the models are believed to be fairly reliable, one could conclude that this is achievable in the reasonably near-term.

5.4 Systematic Model Errors

The major uncertainty for GHG emissions estimates is the extent to which we can believe the transport models developed by the atmospheric science community. This issue becomes exacerbated when schemes that couple car-

bon monitoring to full GCM simulators are proposed, as the issues here are even more uncertain. We will assume instead that meteorological data is being input from separate measurements (or at least that the GCMs are being heavily constrained by data assimilation from widespread measurements) and focus instead on the transport part.

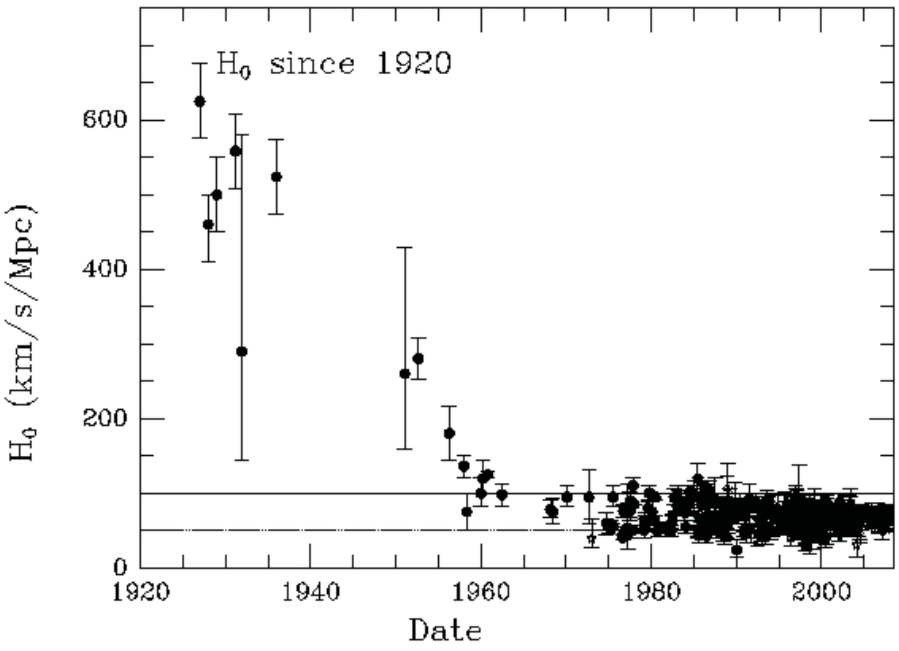


Figure 25: History of values of the Hubble constant. Obviously, early error estimates were totally misleading as to the real nature of the uncertainty, which was not statistical.

The fact that the statistical errors can completely misstate the actual uncertainty in the inferred value of a parameter is, unfortunately, well known. In the physics community, perhaps the most egregious example of error underestimates occurred during attempts to evaluate the Hubble constant relative to the expansion of the universe. (see Figure 25.)

A simple example of how meteorological errors can lead to totally inaccurate fluxes is provided by vertical transport. Assume that a particular measurement system is most sensitive to concentration at a specific height, for example 100 m above the ground. It makes a huge difference, then, if the

model predicts that fluxes from the ground will in fact be advected up to that height or not. Thus, ground sensors in the tropics will underpredict sources if strong vertical mixing is not properly taken into account. One might hope that satellite measurements with that vertical averaging would circumvent this problem. But even this may be wishful thinking. Wind typically exhibits high vertical shear and therefore the extent of horizontal transport is directly coupled to the vertical mixing.

It is absolutely critical to recognize that the goodness of fit of the data by the proposed flux model need not have anything to say about the accuracy of the inferred flux values. Being able to fit the data with a particular model merely relates to the extent to which there are enough flux degrees of freedom and these are all well-determined by either the data or a priori knowledge. In general, there is no reason to expect that one cannot fit data equally well with two different models if one adjusts the fluxes appropriately. This is sometimes investigated in the community by doing the inversion with different models and finding the resulting variation in inferred fluxes. This is a commendable practice. Alternatively, one can perform sensitivity analysis to detect how crucial certain modeling assumptions may be. In any case, modeling errors are the most insidious problem as they will often not give any indication of their presence in the data or the inferred flux values and statistical accuracies.

In the absence of any scheme which will automatically decide if the inferred fluxes are truly insensitive to models, the best approach it to validate schemes by comparison to controlled experiments. In this regard the series of ETEX experiments aimed at detecting the source of trace release was quite informative. In both ETEX-1 and ETEX-2, a variety of advective-diffusive codes were used to invert ground station measurements of PMCP (an inert gas). Meteorological data was taken from standard weather reporting organizations. Surprisingly, the two experiments led to different findings. Whereas in ETEX-1 the inversion was reasonably successful and reasonably independent of modeling details, the inversions of ETEX-2 (see Fig. 26) all vastly

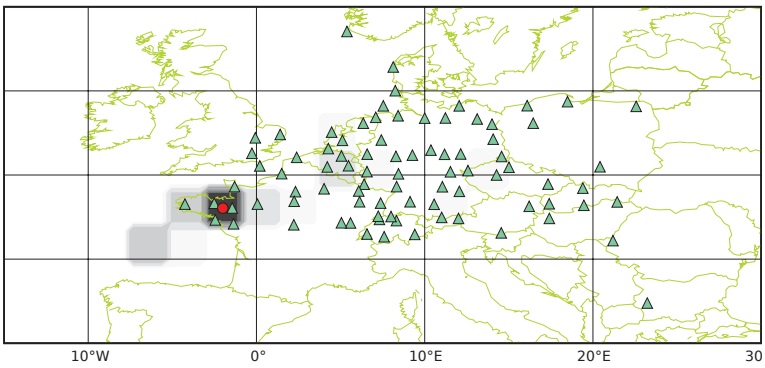


Figure 26: ETEX-II geometry, showing the spatial grid-cell for the resolution of 2.25^0 x 2.25^0. The triangles represent the monitoring stations which gathered the measurements used in the inversion. The disc marks the true release site, Monterfil, France. From [42].

underestimated the source. More than a decade after the experiment there is still no consensus on the source of the difficulty. Results taken from the work of Bocquet and collaborators [42] is shown in Figure 5.4. The source location was accurately determined but the magnitude was way off. One possible reason was measurement error; another had to do with anomalies of large vertical transport (due to unusual meteorological conditions) leading to reduced surface concentrations as compared to what the model had predicted should be the case for the actual experimental flux.

One interesting lesson from the ETEX study is in a general sense consistent with the already mentioned HCFC inversion study of Stohl et al [40]. It appears to be easier in general to infer source locations than absolute values of fluxes. In that study, inventories provided reasonable information about source levels and the inversion algorithms were able to do rather well in assigning sources to specific locations consistent with probable emitting factories. In general, finding source locations depends explicitly on backtracking from an observed concentration plume and the relative ratios of readings from many sites apparently can collaborate so as to give good source locations. But, absolute fluxes cannot be obtained by measurement ratios and really do

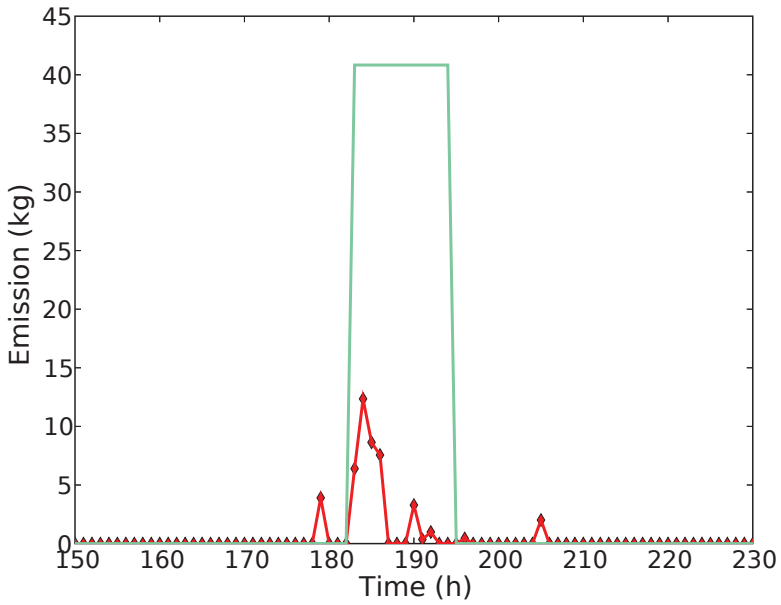


Figure 27: ETEX-II. Profile of the reconstructed source in Monterfil. Total reconstructed mass is equal to 73 kg and the mass reconstructed in Monterfil is 51 kg. Actual release was 490 kg. Spatial resolution is 2.25^0 x 2.25^0. From [42].

require a correct transport model, especially with regard to vertical mixing. This will in general be impossible to get right with any system that has no z-resolved data.

It would obviously be very interesting to do tracer experiments with satellite data. This would help us understand the extent to which the inherent averaging capability of satellite data makes the inversion less sensitive to small-scale modeling errors. Given the importance of controlled release experiments for validating modeling results we recommend that additional new tracer release experiments be conducted. Such experiments and the required capabilities are discussed further in Appendix F and in the section on Findings and Recommendations.

5.5 Reprise

Based on an admittedly short-term look at the state-of-the-art of flux inversion, it is clear that much work needs to be done before we can even come close to claiming that any specific measurement scenario could reach a useful level of true accuracy. This will require much more effort aimed at understanding model limitations, introduction of noise via discretization, correlations between measurements for a given set of flux assumptions etc.

6 SPECIFIC FINDINGS AND RECOMMENDATIONS

Our principal recommendations and assessments can be found in the Executive Summary of this report. In the following section we elaborate on various findings and recommendations related to the topics covered in this study.

6.1 Role of Direct Measurements

Findings:

- Precise determination of a non-cooperative country's annual anthropogenic CO_2 emission ($\equiv F_{Annual}$) through direct CO_2 measurement is difficult due to the following factors:
  - Large spatial and temporal variations in natural background make it difficult to separate out the anthropogenic component from the net emissions.
  - Measurements are typically sparse in both time and space.
  - Model inversion from concentration to flux depends on transport (due to winds and weather) and diffusion, and their uncertainties.
- With planned augmentation of sensors (e.g. the upcoming launch of the Orbiting Carbon Observatory), we judge that reduction of the uncertainty in F_{Annual} to 50% is plausible within about 5 years, to 20% is possible but challenging, and to 10% is unlikely for the foreseeable future.
- Both remote and *in-situ* measurements will be important.
- Methods for determining F_{Annual} can be validated by measurements of U.S. emissions and/or emissions of other cooperative countries, for which independent information about emissions can be obtained.
- The technical goal for improved direct measurement capabilities should be to improve the remote and *in-situ* sampling networks, as well as as-

similated metrology measurements, to the level that estimates of F_{Annual} are limited by natural variability rather than by measurement errors and modeling uncertainty.

Overall Recommendation:

- An organizational unit should be identified or if necessary constituted with the responsibility and authority for developing an overall strategy and systems design for an integrated network of *in-situ* and satellite sensors with the specific goal of optimized direct measurement of GHG emissions from major GHG emitting countries.

We now discuss three specific sub-topics of direct measurements: remote sensing, sensor development and deployment, and isotopic measurements.

6.1.1 Remote sensing

The use of instruments on satellites to measure greenhouse gases is attractive, particularly for the monitoring of emissions from uncooperative countries where *in situ* sampling may not be possible. All currently planned U.S. GHG satellites and instruments are, or will be, in low earth orbit (LEO).

Findings:

- By virtue of their small ground sample distance, LEO satellites will be useful for constraining fluxes from individual sources such as large industrial plants and possibly selected cities. But because of their narrow ground swath and long revisit time, satellites in LEO are not optimal for monitoring anthropogenic emissions from entire states in the face of a natural background that is highly variable in space and time. It is not clear to us that LEO satellites will be able to optimally isolate and measure anthropogenic CO_2 emissions in the face of the large and variable natural signals.

- A geosynchronous (GEO) satellite offers a complementary set of measurements that would be particularly powerful for potential treaty monitoring. In particular it would allow better decoupling of spatial and temporal variability.
- A GEO satellite is technically feasible.

Recommendations:

- We strongly advocate augmenting the planned LEO satellites and instruments with one or more GEO satellites intended for greenhouse gas measurement.
- For the design of these satellites we recommend using model-derived optimization of, e.g., ground sample distance, temporal versus spatial sampling, accuracy, and combining observations of short-wave infrared (SWIR) absorption with observations of thermal IR emissivity. High orbits other than GEO should also be evaluated for the possibility of enhanced performance.

6.1.2 Sensor development and deployment

Several sophisticated measurement techniques for *in situ* CO_2 sensors have been demonstrated, and are available commercially. However for the purposes of monitoring and verification of international agreements, one needs to design a global-scale measurement *system* that would optimize sensor placement and accuracy. Doing so will require tools to address questions such as, "should one use fewer but more accurate sensors, or larger numbers of less-accurate sensors?" The availability of moderate accuracy low-cost sensors will enable a large variety of possible sensor placements.

Findings:

- A large increase in the network of point sensors (ground/air/ships) will very likely be useful for F_{Annual} determination.

- Less costly and more widely deployable point sensors that have suitable accuracy for measuring local CO_2 concentration can likely be developed in the near-term.

Recommendations:

- Pursue the development of a 1 ppm accuracy CO_2 sensor that costs less than $500.
- Use modeling capability to evaluate the value of possible deployments, including considerations of sensor proliferation, location, and accuracy:
  - Cell phone towers world-wide,
  - Weather balloon radiosondes for vertical profiles,
  - Buoys or floaters at sea,
  - Commercial aircraft,
  - Widespread distribution to cooperative sites.

6.1.3 Isotopic measurements

Measurements of CO_2 concentration yield emission estimates on net emissions of CO_2. Information on the anthropogenic sources is derived from the spatial and temporal variations of the CO_2 concentration measurements. Measuring isotopic ratios provides a different approach for determining the processes that produced the CO_2.

Findings:

- Dilution of ^{14}C is a powerful tracer of fossil carbon.
- Complex environmental exchange processes, many of which enrich or dillute abundances of carbon and of oxygen isotopes, complicate quantitative interpretion of isotopic data.
- CO_2 has a rich spectrum of isotopologues (two stable and one long-lived unstable carbon isotope, three stable oxygen isotopes) whose potential has not been exploited.
- There will be a continuing need for facilities for mass-production analyses of isotopic abundances, both unstable (accelerator mass spectrometry) and stable (conventional mass spectrometry and vibration/rotation spectroscopy for remote sensing).
- A baseline library of CO_2 samples from locations of interest would be valuable.
- Isotopic measurements, especially of plant samples, provide time-averaged constraints on fossil-fuel emissions that are complementary to time-resolved direct measurements of CO_2 concentration.

Recommendations:

- Ensure that the US atomic mass spectrometer (AMS) capability be maintained, either by support of the LLNL facility, or by investment in another, comparable AMS dedicated full time to carbon isotope analysis.
- Undertake an annual survey of $^{14}C/^{12}C$ in the U.S. via samples of crops and other vegetation, and validate the inferred mixing ratios of fossil fuel CO_2 against other measures (especially inventories and direct measurements of combustion products such as CO, NO_x).
- Start a pilot program of ^{14}C analysis using crops of known provenance (region, growing season) imported from other nations.

- Create a funding opportunity for research into how spatial and temporal variations in $^{47}CO_2$ can inform understanding about the combustion or other processes that produce it.

6.2 Role of Proxy Observations: Monitoring Energy System Infrastructure

Findings:

Proxy signatures of energy system infrastructure "activities" could include:

- Imagery of the construction and operation of wind farms, solar power generation, and nuclear plants.
- Imagery that samples the transport of coal by rail from the mines to coal-burning power plants, of oil from wells to refineries and of product from refinery to end-user.
- Thermal signatures of crude oil refining distillation columns, of compressors for carbon capture and storage (CCS), and of compressors on CH_4 distribution pipelines.
- Imagery of feedstock flows and sizes of fermenters in biorefineries.
- Other signatures and signals observed by commercial, civil and other systems.

Additional sources of CO_2 emission that could eventually be the object of "proxy" monitoring include iron and steel production, cement production, and activity at industries using fossil-fuels as a feedstock. In the near term monitoring the items listed above would account for a large fraction of the anthropogenic CO_2 emissions from the countries that emit more than 80% of the CO_2.

Recommendations:

- Acquire and maintain a detailed knowledge of the energy infrastructure of the major greenhouse gas emitting countries and identify key signatures of energy flows. To accomplish this, the following actions are required:
  - First, a systems-level strategy must be developed. This strategy should lay out clear long-term goals (e.g., to monitor several progressively more strict limitations on greenhouse gas emissions, with the associated required uncertainties), and determine which countries will be monitored (e.g., the approximately 10 countries responsible today for more than 80% of the greenhouse gas emissions).
  - An organization should be identified or if necessary constituted that will be responsible for putting in place this proxy monitoring system from a combination of civil, commercial, and other resources. This organization, which would be staffed by technical experts from a variety of agencies, would be tasked to determine what activities will be monitored, and which remote-sensing (and possibly *in situ*) assets could best be used for monitoring each activity. They would then use modeling to develop quantitative sampling strategies needed to achieve the necessary levels of uncertainty for each activity. We estimate that a sizeable technical organization will be needed for these tasks over the next 5 years. We also note that pertinent expertise exists within the intelligence community, the Department of Energy (e.g. EIA and Critical Infrastructure groups), as well as other agencies.
  - Begin to calibrate infrastructure-based monitoring signals against known "ground truth" (in the U.S. and other cooperative countries) as soon as possible, once the above goals and strategies have been agreed upon.
  - Begin soon to acquire data from other countries to establish a baseline for comparison with future data.

6.3 Role of Modeling

Large-scale computational models of greenhouse gas emissions and their effect on global climate have developed rapidly over the past decade. However, the main thrust of these models has not been on treaty monitoring or verification. Rather, they have been built and optimized to understand the *science* of climate change, and to assess the *global* effects of greenhouse gases on future climate.

Findings:

- Large-scale climate simulation models have four essential roles:
  - Inversion of concentration measurements to obtain estimates of fluxes from natural and anthropogenic sources,
  - Estimation of uncertainties in the determination of these fluxes, given the spatial/temporal density of sampling, the inherent measurement error of the instruments, and errors in the meteorological models/data used in the conversion of concentration to flux,
  - Distinguishing systematic bias and drift in measurements from actual changes in flux,
  - Identification of the necessary and sufficient measurements needed to estimate emissions (F_{Annual}), and optimization of the sampling strategy for a coordinated system of measurements.
- The last three roles are under-developed, but are critical for optimizing a measurement network and for support of treaty monitoring.

Recommendations:

- Develop the capability to make quantitative estimates of flux uncertainty using simulation models. This should be pursued cooperatively with other relevant programs such as those of NOAA and those carried out in the context of CTBT monitoring.

- Use these models to quantify the sensitivity of flux uncertainty to systematics in measurements.
- Develop modeling tools that can be easily used for design and evaluation of *in-situ* sensor networks and of satellite systems and sampling requirements. Design of *in-situ* networks may likely be country-specific.
- Develop the capability to carry out experiments with controlled emissions of tracer gases in order to validate modeling and inversion tools. This should be done in cooperation with other relevant efforts such as CTBT monitoring.

We remark that significant expertise resides in the academic research community and this expertise should be utilized. However, there needs to be a designated operational entity that has the responsibility for developing, maintaining and maturing the necessary tools for modeling and design.

6.4 Roadmap

Our principal recommendations are summarized in the form of a roadmap provided in the Executive Summary of this report.

References

[1] Keeling, C.D. et al.,*Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980*, *Nature*, **375**, 666-670 (1995).

[2] *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements*, National Research Council, ISBN: 0-309-15212-7 (2010).

[3] *Climate Change 2007: Mitigation of Climate Change. Synthesis Report*, Core Writing Team, Pachauri, R.K. and Reisinger, A. (eds.), IPCC, Geneva, Switzerland (2007).

[4] http://unfccc.int/kyoto_protocol/items/2830.php

[5] http://unfccc.int/home/items/5262.php

[6] *Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (2007)

[7] http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/ and www.carbon tracker.eu

[8] http://www.ramgen.com/apps_comp_storage.html

[9] http://www.eia.doe.gov/cabs/China/Full.html

[10] http://www.eia.gov/oil_gas/petroleum/data_publications/ refinery_capacity_data/refcap_historical.html

[11] K.R. Gurney, D.L. Mendoza, Y. Zhou, M.L. Fischer, C.C. Miller, S. Geethakumar, and S. de la rue du Can, *High Resolution Fossil Fuel Combustion CO_2 Emission Fluxes for the United States'*, Environ. Sci. Technol., 43 (14), pp 55355541 (2009).

[12] Seminar presentation at the NASA Jet Propulsion Laboratory, September 2010.

[13] http://www.legos.obs-mip.fr/en/equipes/gohs/resultats/d_alti_lacscont

[14] Nondestructive Testing Handbook, 3rd Edition, Vol 3, Infrared and Thermal Testing, Chapter 16 (2001).
(http://www.asnt.org/shop/143wcd/ir16.pdf)

[15] Rayner, P. J., M. R. Raupach, M. Paget, P. Peylin, and E. Koffi (2010), *A new global gridded data set of CO_2 emissions from fossil fuel combustion: Methodology and evaluation'*, J. Geophys. Res., doi:10.1029/2009JD013439, in press. (accepted 27 April 2010).

[16] http://www.purdue.edu/eas/carbon/vulcan/index.php

[17] http://www.bwea.com/aviation/ams_report.html

[18] H. Bovensmann, M. Buchwitz, et al, *A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications* Atmos. Meas. Tech., 3, 781-811 (2010).

[19] Z. Kuang, et al, *Spaceborne measurements of atmospheric CO_2 by high-resolution NIR spectrometry of reflected sunlight: an introductory study* Geophys. Res. Lett. 29, 1716 (2002).

[20] C. E. Miller, et al, *Precision requirements for space-based X_{CO2} data* J. Geophys. Res., 112 (2007).

[21] M. Chahine et al, *On the determination of atmospheric minor gases by the method of vanishing partial derivatives with application to CO_2.* Geophys. Res. Lett., 33, L038 (2005).

[22] D. Crisp et al., *The orbiting carbon observatory (OCO) mission*, Adv Space Res, 34, 700-709 (2004).

[23] H. Matsueda et al, *Aircraft observation of carbon dioxide at 8-13 km altitude over the western Pacific from 1993 to 1999* Tellus, 54B, 1-21 (2002).

[24] Riley Duren, private communication (2010).

[25] Guilderson, T. P., briefing to Jason, June (2010).

[26] Burns, S.P., *et al.*, *An Evaluation of Calibration Techniques for In Situ Carbon Dioxide Measurements Using a Programmable Portable Trace-*

Gas Measuring System, Jour. of Atmos. and Oceanic Tech., **36**, 291 (2009).

[27] Burrows, J.P., *et al.*, *The geostationary tropospheric pollution explorer (GeoTROPE) mission: objectives, requirements and mission concept*, Adv in Space Research **34**, 682 (2004).

[28] Olsen, S.C. and Randerson, J.T., *Differences Between Surface and Column Atmospheric CO_2 and Implications for Carbon Cycle Research*, JGRD 10902301O (2004).

[29] Stephens, B.B., Watt, A., Maclean, G. (2006), *An autonomous inexpensive robust CO_2 analyzer (AIRCOA)*. 13th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques, J. Miller ed., WMO TD no. **1359**, 95-99. (2006).

[30] www.eol.ucar.edu/stephens/RACCOON/ .

[31] Orphal, J. *et al.*, *Monitoring tropospheric pollution using infrared spectroscopy from geostationary orbit*, C.R. Physique **6**, 888 (2005).

[32] L Rothman et al, JQSRT 110, 533-572 (2009).

[33] Silver, J.A., and Zondlo, M.A, Proc. SPIE Vol 6378, 63780J-1 (2006).

[34] *A single gas chromatograph for accurate atmospheric mixing ratio measurements of CO_2, CH_4, N_2O, SF_6 and CO*, S. van der Laan, R. E. M. Neubert, and H. A. J. Meijer, Atmos. Meas. Tech. Discuss., 2, 13211349 (2009).

[35] *Comparison between real time and flask measurements of atmospheric O_2 and CO_2 performed at the High Altitude Research Station Jungfraujoch, Switzerland*, Chiara Uglietti, Markus Leuenberger, and Francesco L. Valentino, Science of The Total Environment Volume 391, Issues 2-3, 1 March 2008, Pages 196-202 (2008).

[36] www-pm.larc.nasa.gov/triana/NAS.Triana.report.12.99.pdf

[37] Djuricin, S., Pataki, D. E., and Xu, X., *A comparison of tracer methods for quantifying CO_2 sources in an urban region*, JGR, 115, D11303 (2010).

[38] G. Toon, J. Blavier, R. Washenfelder, D. Wunch, G. Keppel-Aleks, P. Wennberg, B. Connor, V. Sherlock, D. Griffith, N. Deutscher, and J. Notholt, *Total Column Carbon Observing Network (TCCON)*, in Hyperspectral Imaging and Sensing of the Environment, OSA Technical Digest (CD) (Optical Society of America, 2009), paper JMA3.

[39] *Review of Scientific Aspects of the NASA Triana Mission: Letter Report*, National Academies Press (2000).

[40] A. Stohl, J. Kim2, S. Li, S. O'Doherty, J. Muhle, P. K. Salameh, T. Saito, M. K. Vollmer, D. Wan, R. F.Weiss, B. Yao, Y. Yokouchi, and L. X. Zhou, *Hydrochlorofluorocarbon and hydrofluorocarbon emissions in East Asia determined by inverse modeling*, *Atmos. Chem. Phys.* **10**, 3545-60 (2010).

[41] P. J. Rayner and D, M. O'Brien, *The utility of remotely sensed CO_2 concentration data in surface source inversions*, *Geoph. Res. Lett.* **28**, 175-8 (2001)

[42] M. Krysta1, M. Bocquet, and J. Brandt, *Probing ETEX-II data set with inverse modeling*, *Atmos. Chem. Phys.* **8**, 3963-71 (2008).

[43] F. Chevallier, F.-M. Breon, and P. J. Rayner, *Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework*, *J. Geophys. Res.* **112**, D09307 (2007).

[44] M. C. Bocquet, *Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: Theory*, *Quat. J. Royal Meteorol. Soc.* **131**, 2191-2208 (2005).

[45] P. Bergamaschi, C. Frankenberg, J. F. Meirink, M. Krol, M. Gabriella Villani, S. Houweling, F. Dentener, E. J. Dlugokencky, J. B. Miller, L. V. Gatti, A. Engel, and I. Levin, *Inverse modeling of global and regional CH_4 emissions using SCIAMACHY satellite retrievals*, *J. Geophys. Res.* **114**, D22301 (2009).

A APPENDIX: JASON and Inexpensive NDIR

We purchased three SenseAir instruments from CO2Meter in order to evaluate the opportunities for achieving high accuracy by appropriate experimental use. Figure 28 illustrates one of these boards, and other equipment we assembled. The K30 "sensor development kits" we purchased cost $200

Figure 28: The SenseAir sensor is shown on the left, a small 6V air pump purchased from CO2Meter, a desiccant (white) and CO_2 scrubber (black) we created by filling an aquarium check valve with CaO, our P and T datalogger from GCDataConcepts is on the right, and a CO_2 cylinder is below. The CaO scrubber can remove the CO_2 from about 500 liters of air and the CO_2 cylinder can fill about 8 liters at STP.

apiece, but the boards are sold for $112.50 in quantity 10 or more. We also purchased from CO2Meter a small, battery-driven pump, desiccant, and sensor cover. We obtained bottles of CO_2 intended for bicycle tire inflation, and we also removed the valve portion of an aquarium check valve and inserted about 5 grams of CaO wrapped in gauze into the volume to act as a CO_2 scrubber.

CO2Meter also distributes another sensor from Gas Sensor Systems, the C100, that uses two LEDs and photodiodes, tuned to the CO_2 band at 4.26 μm and an off-band at 3.95 μm. We have no reason to believe that

this sensor might not have superior performance to the K30, but our goal was to evaluate the performance of one exemplar rather than trying to do an industry-wide comparison of all the options.

The SenseAir K30 board uses pyroelectric sensors that detect flashes of incandescent light through a narrow band filter every 2 sec. They are quite sensitive to temperature, so the detector uses a "blind channel", a matched detector that is masked, and the difference of these two raw signals provides the signal S. The light source is carefully qualified and the drive signal for it is regulated. The light is emitted into a gold-plated labyrinth, so although the total size is about 3 cm, the path length L is about 10 cm. SenseAir claims that their "curve of growth" calibration is stable at the few percent level, but that gain drift limits their accuracy to ± 30 ppm. (The use of a differential, blind channel eliminates zero point offset to some accuracy.) The unit is designed to work to about 1% CO_2, so detection of 400 ppm (0.04%) is at the very low end of the designed operation range. (SenseAir somewhat confusingly calls their gain calibration the "zero" offset, because it is the response with zero concentration CO_2.)

Pyroelectric sensors use the heating from illumination to deform a crystal and generate a dipole separation of charge which can be detected as a momentary voltage. They therefore have no DC response at all, and their use requires a variable light signal such as the flashing incandescent light. They cannot provide a signal at zero light intensity, and so it is not possible to obtain an offset measurement this way, hence the use of a blind channel as reference. In fact the outputs from the lit and blind channel vary together dramatically, and taking their difference reduces the variation RMS by a factor of 15. Photoelectric detectors such as PbSe, InGaAs, or InSb do offer the opportunity to measure a response at zero illumination.

The K30 sensor employs two software features to improve the utility of the output. The first is "automatic baseline correction" (ABC), that adjusts the lowest value baseline output back to 400 ppm with an adjustable time constant, 7.5 days by default. The second is "adaptive leaky averaging" of

the output, an exponential filter with an e-folding that changes between 0 and 256 samples depending on the perceived rate of change. For all experiments we disabled the ABC, but left the leaky average function active.

The observed intensity in both lit and blind channels is available from the USB interface, as well as the difference and a filtered intensity difference. This filtered intensity is converted to a "calibrated intensity" according to the gain factor ("zero adjustment" in SenseAir's terminology), and then to a CO_2 concentration via linear interpolation between a 16 point table of intensity, CO_2 concentration pairs that is determined for each unit and stored in EEPROM.

We also purchased a pressure and temperature data-logger from Gulf Coast Data Concepts (gcdataconcepts.com), knowing that the indicated concentrations would depend on P and T. This data-logger measures pressure and temperature to accuracy of about 0.1%, runs for about two weeks on an AA battery, records to a 1GB microSD card, and has a USB plug for computer communications. It costs $119. There are also many USB humidity sensors available, and we believe that incorporation of pressure, temperature, and humidity sensors onto a board for CO_2 measurement would involve a cost of $\sim$\$30, and negligible power, mass, and size.

Our first step was to write software to communicate with the K30 board in order to exercise all visible functions and extract data more flexibly than the supplied Windows GUI. Very helpful discussions with CO2Meter revealed access to many of the functions that are not available through the FTDI interface, such as reading of intensity, examination of RAM and EEPROM, and disabling the leaky averaging or ABC calibration.

The three boards report CO_2 concentrations every 2 seconds, and with the ABC and leaky averaging disabled the CO_2 values have distributions that are reasonably independent and Gaussian. The RMS of reported CO_2 is 16 ppm, 9 ppm, and 10 ppm for the three boards respectively, and 2.8 ppm, 1.4 ppm, and 2.2 ppm when averaged over a 100 sec interval and drift removed. With

the leaky averaging enabled (effectively averaging some 256 prior samples) the RMS is about 1.5–2 ppm.

Our first experiment began on June 30, when we dangled three SenseAir units outside of the second story window of our office building. The site is on a campus that has managed, California vegetation, consisting mostly of eucalyptus trees and grass. There are many parking lots nearby with occasional traffic, the I5 freeway is about 0.5 km east, the I805 freeway is about 2 km east, and roads that carry substantial traffic are also within 1 km. Breezes were light and we did not notice any particular prevailing direction. Figure 29 illustrates what we observed from one of the sensors over a period of two weeks.

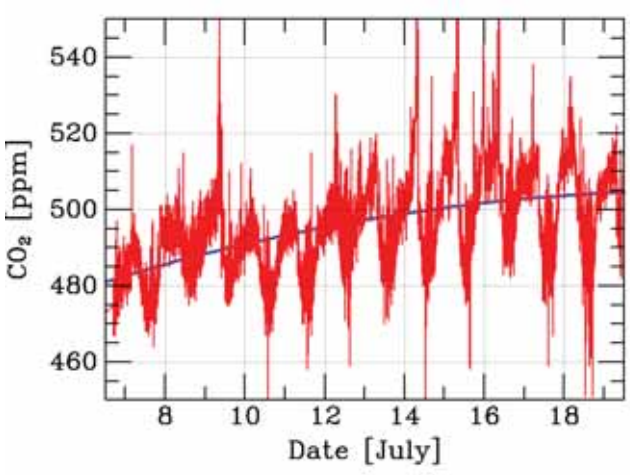


Figure 29: Two weeks of CO_2 concentration outside Building 29 is shown for one of the sensors. The drift indicated by the blue line is probably a sensor artifact; the shorter term variation is real. July 10, 11, and 17,18 were weekends.

The sensors were all set to 400 ppm on June 30, and we believe that the steady upward drift in reported concentration is instrumental drift. These units are not supposed to attain full stability until a 3 week "burn-in" period, so the roll-over in mean seen in Figure 29 may indicate the ultimate stability is being achieved. We have seen less drift since this period, but we believe

that a 10 ppm/day drift can be expected, and a 1 ppm precision with a 400 sample, 800 sec integration time.

Despite the slow drifts in indicated CO_2 concentration, we were immediately struck by the dramatic variability, both episodic and diurnal. Figure 30 shows how well correlated these are between sensors. The correlation of CO_2

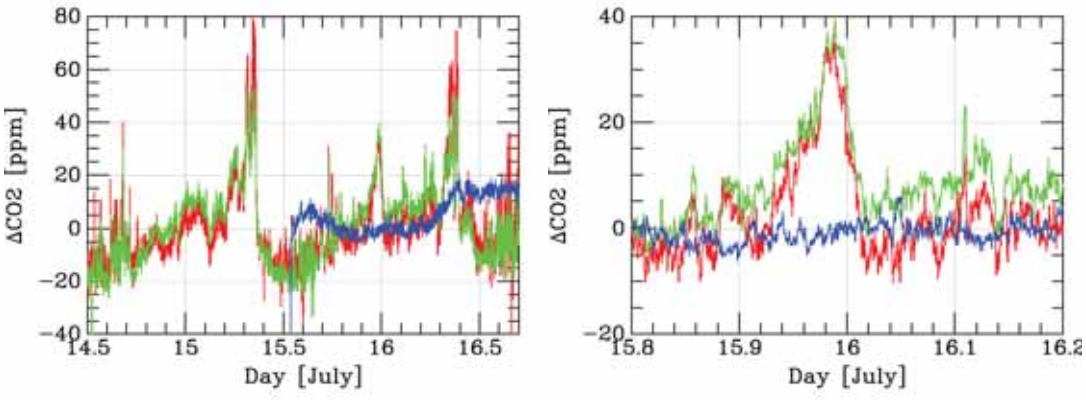


Figure 30: Expanded time history of CO_2 outside Building 29. The three sensors are shown in different colors. The "blue" sensor was placed in a plastic bottle whose slow leakage provided a ∼8 hour time filter on the CO_2 concentration seen. The RMS variation of the blue samples in the right hand panel is 1.9 ppm.

spikes with CO_2 emissive events such as leaf blowing, grass cutting, and tree trimming, and also the abrupt diminution of spike frequency at 6 pm each day convinces us that they are real. These spikes have a very large amplitude compared to the accuracy of the SenseAir boards, and clearly convey a great deal of information. Whether this information is useful remains unclear, however. There is a significant diurnal variation that is surely caused by plant respiration, most clearly visible on the weekends, although the onset of beach traffic obscures the dip in CO_2 in the middle of the day.

We did think of separating the three sensors so as to cross-correlate the reported concentrations as a function of time and be able to determine the size and direction of the CO_2 clouds blowing by, but did not have an opportunity

to perform that experiment.

In order to obtain an absolute measurement of CO_2 concentration from the SenseAir unit, we must find a way to calibrate its systematic bias and drift. Figure 31 shows the basic idea. The net transmission through the cell follows

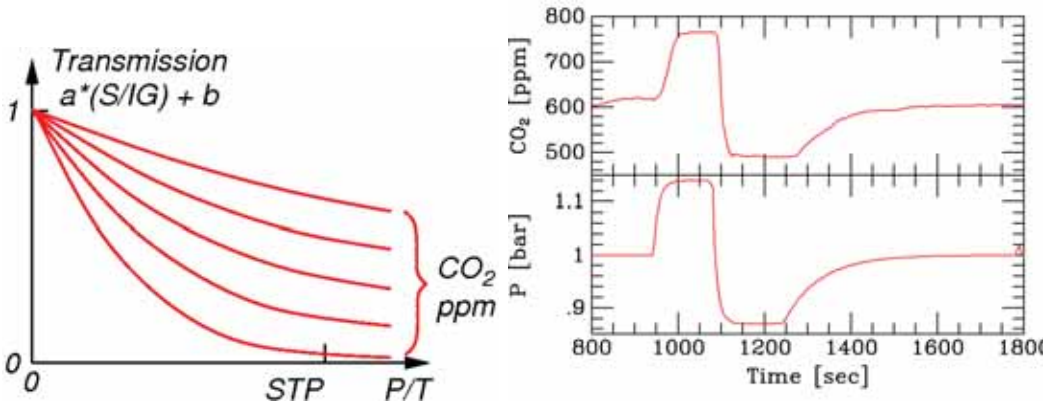


Figure 31: A notional graph showing the effect on transmission through a fixed gas cell on CO_2 density (proportional to P/T) and on CO_2 concentration is shown on the left. The transmission should be proportional to the measured signal S divided by the light intensity I and system gain G, with some offset and scale to be determined. The effects we observed by varying P are shown on the right.

a set of curve of growth curves that depend on the number density of CO_2 molecules, and therefore on the PC/T combination in Equation 4-3. In the face of systematic drifts, we need three pieces of information to determine a concentration, the instrumental offset, and gain. For the SenseAir the offset is supposed to be very small because of the matched, differential input between lit and blind channels.

The ideal, commonly used for calibration of the instrumentation at CO_2 monitoring sites, is to measure output for gas samples at three, known concentrations that straddle the concentration of interest. (The non-linearity of a curve of growth is substantial enough that a linear interpolation between two samples is likely to be inaccurate unless the gas samples happen to be very close to the unknown concentration.) For example, at the Mauna Loa

observatory, three reference gas bottles of concentration calibrated at the $\sim$0.07 ppm level from the WMO are used, at concentrations of about -10, 0, and $+10$ ppm relative to ambient (Tans and Thoning, 2008).

CO2Meter recommends taking the curve of growth of the SenseAir as fixed and known, calibrated at the factory into EEPROM, and measuring only the gain. They provide the means to adjust this gain to give the correct concentration at "zero CO_2" by exposure to gas at 0 ppm or 400 ppm concentration. (Their ABC algorithm automatically shifts the offset to make the 7.5 day lowest concentration 400 ppm.) We investigated three methods of calibration.

The right panel of Figure 31 shows the effect of changing pressure on the CO2Meter's indicated CO_2 concentration. We placed the sensor in an air tight bottle and alternately connected the small pump we purchased to pressurize or evacuate the jar. We monitored the pressure in the jar using the data-logger. This tiny pump is evidently capable of changing pressure by about 12% relative to STP, and we observed the expected change in the indicated concentration. In principle, we are measuring the slope of the curve in Figure 31 at STP, and this can be one of the two measurements needed to set the offset and gain.

A more direct method of calibration is shown in Figure 32. We placed within the jar approximately 5 g of CaO wrapped in filter paper which has a voracious appetite for CO_2. Even without air circulation, the CO_2 concentration dropped with a time constant of about 400 sec, effectively reaching 0 ppm within an hour. This is another independent measurement of offset and gain. If we take the offset as fixed and zero, this alone is sufficient to calibrate the SenseAir unit. We did not have the means to evaluate the accuracy of subsequent concentration measurements of ambient air at 400 ppm (i.e. verify the SenseAir calibration of the curve of growth).

The last calibration method involves trying to create a known concentration by mixing known volumes of gas. This is illustrated in Figure 32. We mea-

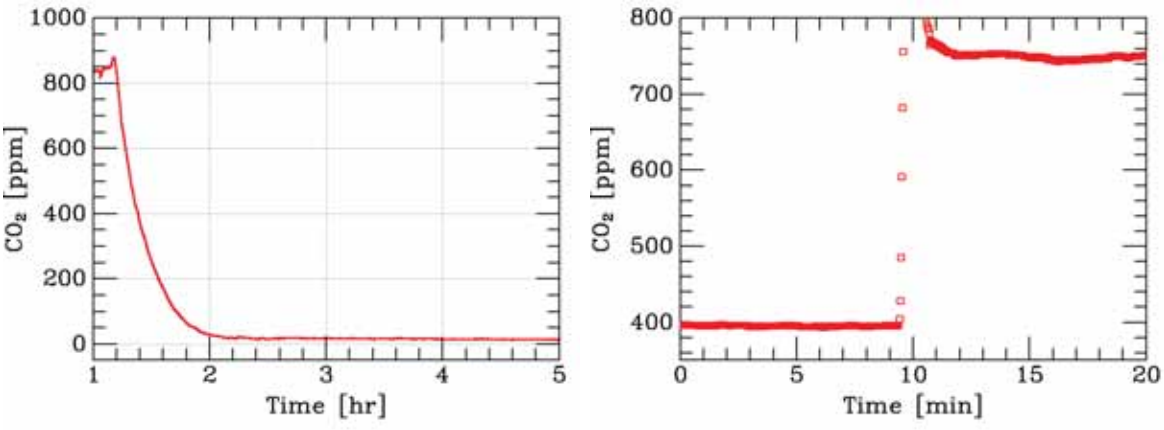


Figure 32: The result of exposing the sensor in a closed bottle with ambient air to a CaO CO_2 scrubber is shown on the left. On the right, a known volume of CO_2 injected into the bottle causes the indicated concentration to rise by 400 ± 47 ppm. The overshoot and subsequent oscillation are artifacts of the dynamic leaky averaging; the RMS prior to injection was 0.9 ppm; the RMS after was 2.8 ppm.

sured the volume of the closed bottle to be 500 ± 30 ml, and we filled another bottle with "pure" CO_2 from a bicycle inflation cylinder. We do not know the resulting CO_2 concentration, but we believe it is probably very close to 100% just by virtue of the relatively low volume and pressure in the cylinder and the care with which we removed the air before filling with CO_2. We then filled a small syringe with CO_2 and injected a volume of 0.2 ± 0.02 ml into the bottle. We therefore expected the indicated CO_2 concentration to rise by 400 ± 47 ppm. In fact the increase in concentration was 350 ppm, indicating that the gain of the SenseAir unit is consistent with being accurate. Obviously this was a crude test, and we started with gas at ambient 400 ppm instead of gas that had been scrubbed of CO_2 but it illustrates the potential of creating known CO_2 concentrations cheaply, on site.

We close with thoughts on how to build a 1 ppm CO_2 sensor for less than \$500. (In fact we believe this could cost no more than \$200 in quantity.) Figure 33 shows the basic idea.

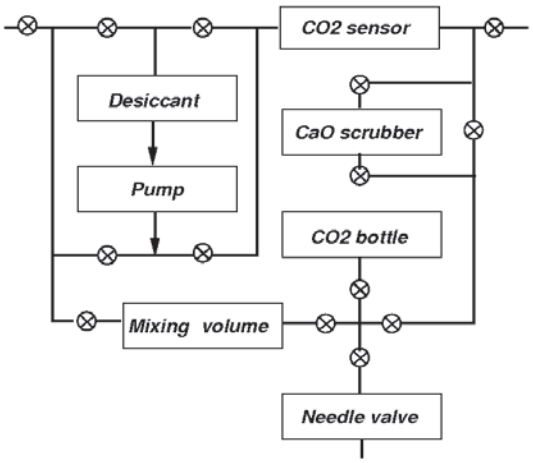


Figure 33: Our concept for a SenseAir equipped with auxiliary equipment capable of providing an absolute calibration. The various valves can be set to feed the sensor ambient air, air scrubbed of all CO_2, air with a measured concentration of CO_2, and air that has been lowered or raised in pressure by 10%.

The heart of the unit is the CO_2 sensor, notionally the SenseAir K30, but others would work as well. Not shown are sensors for pressure, temperature, and humidity. These are all widely available as full, USB-interfaced units for low cost, so we believe the CO_2 sensor board could be equipped with the added functionality for no more than $30.

There are 14 valves shown, but actually there are only a few functions so many of the valve functions can be ganged into a special, plastic piece, for example o-rings drawn along a cylinder, connecting different chambers to input and output ports at different positions. The four valves around the pump and desiccant, for example, simply serve to reverse the flow of the pump, and could potentially be replaced with a DC motor and bidirectional pump.

The five functions provided are

- Direct flow of ambient air drawn through desiccant across the CO_2 sensor for measurement of ambient CO_2 concentration.

- Increase of the pressure on the CO_2 sensor
- Decrease of the pressure on the CO_2 sensor
- Closed cycling of air through a mixing volume and CaO scrubber, producing a 0 ppm atmosphere for the CO_2 sensor with some time constant, and filling the mixing volume with 0 ppm air.
- Creation of a known, calibration concentration of CO_2, by a) flushing the cross shaped volume with pure CO_2 from the bottle and out the needle valve, b) letting that volume come to ambient temperature and pressure, c) mixing this volume of CO_2 into a previously scrubbed volume in the mixing volume (and other plumbing).

We believe it should be possible to achieve 0.1% repeatability with equipment of this sort, and accuracy by a one-time calibration. Note that 0.1 mole of CaO (5.6 g), and a 16 g (8 l at STP) bottle of CO_2 weigh less than an ounce and are adequate for about 1000 calibration cycles.

The SenseAir K30 is designed to measure up to 1% CO_2, i.e. 10,000 ppm. The labyrinth through which the 4.26 μm light passes is about 3 cm in size, and is designed to have a transmission of 93% at 400 ppm CO_2 and 64% at 10,000 ppm. Comparison with Figure 15 suggests that it apparently has an effective path length L of $\sim$10 cm, but the filter is somewhat broader than 2300–2400 cm^{-}1. Therefore at $\sim$400 ppm the attenuation is quite small, placing stringent accuracy requirements on the light source, sensor, and electronics to measure concentrations at the 1 ppm level. That the K30 does as well as it does is therefore quite impressive.

It seems straightforward to design a better labyrinth that has a longer effective path length for a similar compact size. For example, an integrating sphere of diameter R with an input and exit port of size r will emit a fraction of light from the exit port

$$T = \frac{1}{2}\left[1 + \frac{2R^2}{Ar^2} + \frac{2\pi R^3 \alpha}{3r^2}\right]^{-1}, \qquad \text{(A-10)}$$

where A is the absorption per bounce off of the wall and α is the absorption

per unit length, $\alpha(\lambda) = n\,\sigma(\lambda)$ from Equation 4-2 (in the limit that $A \ll 1$, $\alpha R \ll 1$, and assuming random scattering from the walls).

Since the reflectivity of gold at 4 μm is $\sim$99%, we could achieve an effective path length 25 times greater than the diameter of a sphere by using a pair of ports that are 5 times smaller than the sphere. When we integrate this over a bandpass of 4.2 μm to 4.35 μm with notional parameters for velocity and pressure broadening, for example, we find a 25 mm integrating sphere with a pair of 5 mm ports and wall reflectivity of 98% will suffer an extinction of output by 33% when 400 ppm CO_2 gas is present. Reference to Figure 15 indicates that this little 1 inch integrating sphere is therefore acting like a 1 meter path length.

Changing pressure or temperature affects the absorption both through the density of scatters as well as the pressure-broadened line profiles, so we again have a non-trivial curve of growth. Inasmuch as exact details of any inexpensive sensor such as bandpasses, detector and emitter responses, and exact assembly geometry cannot be controlled with any precision, we expect such a sensor would require a calibration process that spans the range of concentration, pressure, and temperature that is expected to be encountered. The *stability* should be possible to control at the 0.1% level, however. If extreme change in pressure is expected, as with a balloon flight to the stratosphere, a differential comparison with a known reference at the same pressure, akin to the design of Silver and Zondlo (2006) might be appropriate.

B APPENDIX: Aircraft Monitoring of Carbon

In monitoring nations' activities in mitigating—or worsening—the global burden of atmospheric carbon, aircraft have a unique role, because they can take vertical profiles of CO_2 and other greenhouse gases from low altitudes above ground-based samplers throughout a column reaching to the top of the troposphere, at selected locations. Measuring carbon fluxes from the ground, or even tall towers, may not sufficiently enable models to identify fluxes from concentrations, because there are appreciable gradients in carbon concentration as a function of altitude in the troposphere. Satellite measurements, although (as we discuss elsewhere) may be of great use in a general carbon-activity monitoring framework, cannot make the detailed vertical profile measurements, including air sampling, that aircraft can make. Nor can they be guaranteed to be where they are needed at a specific time. Aircraft, both commercial with limited instrument packages, and others dedicated to CO_2 monitoring and more fully equipped, are valuable in many respects, as we discuss here.

There are three classes of aircraft programs: The first is already in operation, to some extent, with MOZAIC and similar agreements calling for commercial passenger jets to carry instrumentation that sample vertical profiles at takeoff and landing in major cities. The second consists of aircraft flying dedicated science missions in and under the sponsorship of various nations. The third would be similar in spirit to the Treaty on Open Skies that came into effect in 2002. This agreement, first proposed in the Eisenhower administration, monitors nuclear weapons-related activities in 34 nations, including major countries such as the US, Russia and certain East European nations, France, Germany, Spain, the United Kingdom, and Ukraine. The Treaty on Open Skies calls for monitoring of the entire territory of a signatory nation with aircraft having a specified complement of sensors, with each such nation obligated to allow over flights according to a quota, and allowing the nations

to conduct as many observation flights as its quota of over flights that it must allow. The reason for this complementary program is that the aircraft flight paths will not be limited to commercial routes; in particular, Open Skies aircraft could fly vertical profiles other than near airports.

In this section, we discuss commercial and selected government programs, including MOZAIC, and a possible regime for a Open Skies treaty on carbon monitoring.

B.1 Airline *in-situ* Measurements

Researchers have indicated that a lack of CO_2 measurements over continents is partially responsible for the large variation in CO_2 flux models [1, 2, 3]. There are a number of previous and current efforts to carry out in-situ CO_2 measurements in aircraft, both domestically and internationally. Previous efforts in *in-situ* aircraft measurements can be separated into two categories. The first category is leveraged commercial aircraft that have instruments installed, the second category consists of aircraft dedicated to measurement. The former category of aircraft do not alter their flight path to accommodate any special type of measurement. Here we will detail the previous efforts and then discuss how these efforts may be improved on.

MOZAIC (1994-2009), an EU effort, was originally designed to assess the effects of air traffic in the upper atmosphere (the boundary of troposphere and the stratosphere) [4]. MOZAIC outfitted five Airbus A340 with a 120 kg instrument rack with instrument sampling intake tubes that breach the airplanes' hull. Since 1994, the MOZAIC program has collected data from over 30,000 flights, and has gone through a number of revisions, adding CO and NO_x measuring capabilities in 2002 and 2001, respectively. Figure 34 shows the MOZAIC instrumentation on an aircraft.

Lufthansa has simultaneously run the CARIBIC program (since 1997) and conducts several test flights a year using an A340-600 aircraft with several

Figure 34: MOZAIC instrumentation panel on the Airbus A340 (figure from Volz-Thomas, A., and Cammas J. P.: From MOZAIC to IAGOS Status and Persepectives, online presentation: http://www.wmo.int/pages/prog/arep/gaw/documents/Volz-Thomas_MOZAIC_IAGOS_GAW2009May09.pdf).

tons of equipment that sample over 40 gases [5]. CARIBIC uses the infrared absorption technique to measure CO_2 concentrations in-situ [5]. Current efforts are underway to develop the successor of MOZAIC, which is named IAGOS. The IAGOS project intends to expand the number of participating aircraft and number of gases measured while still using a MOZAIC-like instrumentation payload (hundreds of kg). These efforts, MOZAIC, CARIBIC, and IAGOS are all conducted on commercial aircraft (principally the Airbus A340) and in conjunction with commercial airlines (Lufthansa, Air France, etc.). Figure 35 gives some details.

In addition to the category-1 type measurement craft, the EU had two additional pilot programs (CAATER-1 and CAATER-2) that used DLR-Falcon 20 aircraft for dedicated measurement of CO, CO_2, and 222Rn (Xueref-Remy

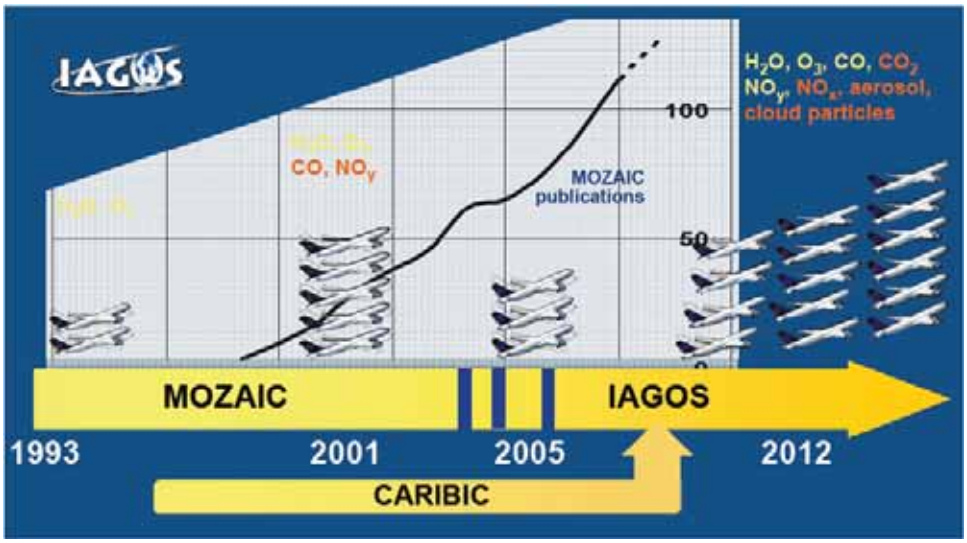


Figure 35: MOZAIC, CARIBIC, and IAGOS timeline, participation, and number of publications ((figure from Volz-Thomas, A., and Cammas J. P.: From MOZAIC to IAGOS Status and Perspectives, online presentation: http://www.wmo.int/pages/prog/arep/gaw/documents/Volz-Thomas_MOZAIC_IAGOS_GAW2009May09.pdf))

2010). These aircraft measured CO_2 via non-dispersive infrared gas analysis and collected samples for in-lab verification. The CAATER-1 campaign consisted of five flights and 14 hours of air time, while the CAATER-2 consisted of three flights and 8 hours of air time with vertical profiles collected between ground and 4000m. Three-dimensional CO_2 profiles from CAATER-1 are shown in Figure 36.

In the second category, dedicated science flights, we note that the USA has had a NOAA program (since 1992) that falls into category-2 aircraft measurements. The Carbon Cycle Greenhouse Gases (CCGG) group uses Cessna 402 aircraft to measure CO_2 and other gases via sample collection at up to 8000 meters. These samples are then measured in the laboratory.

The modeling community might benefit greatly from the implementation of further category-1 airplane measurement programs. At a mass of $\sim$100 kg for instrument packages, the real cost to the airline is a substantial fraction

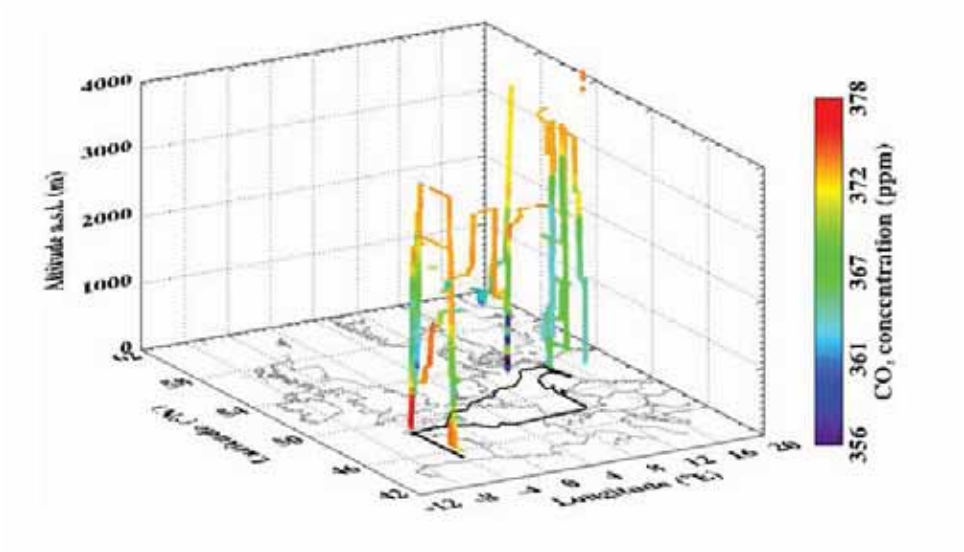


Figure 36: Three-dimensional CO2 profiles from the CAATER-1 program (Figure from Xueref-Remy, I., Bousquet, P., Carouge, C., Rivier, L., Viovy, N., and Ciais, P.: Variability and 20 budget of CO2 in Europe: analysis of the CAATER airborne campaigns Part 2: Comparison of CO2 vertical variability and fluxes from observations and a modeling framework, Atmos. Chem. Phys. Discuss., 10, 42714304, 2010, http://www.atmos-chem-phys-discuss.net/10/4271/2010/.).

of a passenger ticket price, implying that it is reasonable to allocate as much as $100k for the instrumentation. Any large expansion of measurement programs will either involve military aircraft or require that commercial airlines be paid revenues for carrying the greenhouse gas sensors. Therefore it is sensible that the category-1 programs use relatively sophisticated and expensive instrumentation, and sample gasses other than CO_2 such as H_2O, CO, CH_4, O_3, NO_x, NO_y, etc. With newer NDIR measurement technologies, we can hope to see instrumentation packages decrease in cost, size, and mass. Furthermore, providing the devices with a direct feed to the outside air intake, will obviate the need for physical alteration to the craft (unlike MOZAIC and IAGOS).

Hundreds of thousands of flights left US airports in 2009. Boeing 737 and 757 aircraft were the most widely used type of aircraft with nearly 60,000 flights.

Fitting only these two types of aircraft with sampling instruments and data recording devices could provide a wealth of data. A deployment example of such a program could be with Southwest Airline and The United Parcel Service (UPS), which combined, accounted for over 20,000 flights using 737 and 757 aircraft.

There are presently about 4000 Boeing 737 aircraft, 3000 Airbus A320, and 1000 Boeing 757s in service. For the cost of OCO, it would be possible to equip 1000 aircraft with $100k of gas monitoring equipment and allocate $100k per year per aircraft for two years operations cost. Each aircraft flies some million miles per year, so 1000 aircraft would cover 1.6 billion kilometers. By contrast OCO orbits at ∼7 km/s and therefore traverses only 0.2 billion kilometers in a year, only a tenth of the path. Aircraft cannot fly into uncooperative airspaces; aircraft measurements may be well or poorly matched for important flux localization and estimation. Figure 37 illustrates the degree of global coverage possible. Whether this is an opportunity worth

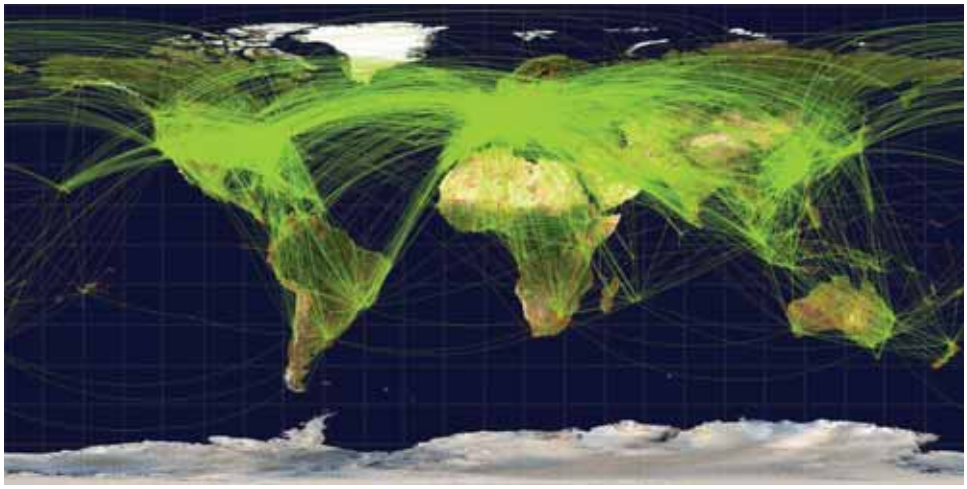


Figure 37: The routes followed by aircraft over the course of a year are shown. (`http://en.wikipedia.org/wiki/ File:World-airline-routemap-2009.png`.)

pursuing depends on how informative it is to models, but this example serves to illustrate how vast the opportunities are, given the vastness of resources the world spends on energy and is willing to spend to maintain a stable climate.

B.2 Open Skies Aircraft Program

The commercial-aircraft programs are very valuable but potentially limited, both because of the nature of the instrument packages they carry and because of limitations on their flight paths illustrated in Fig 37. It is also not clear that their measurements would be accepted for treaty purposes. An alternative would be a fleet of planes whose instrumentation and flights are governed by a treaty. Possible instruments and roles for such treaty aircraft might be

1. Vertical profile sampling in selected tropospheric columns, both near major cities (as to some extent done by MOZAIC, IAGOS,...) and elsewhere (Open Skies), as allowed by treaty.
2. LIDAR profiles from the ground to the treetops through forest canopies too dense for foliage-penetration (FOPEN) radar to be of much use, for monitoring of land use and reforestation.
3. Small, light K-band synthetic aperture radars are capable of high-resolution (down to a few cm, in some cases) imaging, day or night, and can detect change in water, ground, or canopy heights at this high level of resolution if needed. They are also capable of coarser resolution for larger-scale monitoring and change-detection activities.
4. High-resolution EO/IR monitoring of power plants, factories and other man-made carbon-generating activities, if questions of compliance come up and if the resolution is allowed by treaty.

Sampling of vertical atmospheric columns might involve capturing air samples at various altitudes, or using absorbers that allow air to flow through a filter that absorbs carbonaceous compounds. These can be returned to the ground for study, notably for determination of isotopic fractions of ^{14}C, without having to return the entire air sample containing the carbon.

A FOPEN radar typically operates at L-band or longer wavelengths (that is, ≥ 30 cm), in order to penetrate canopies and observe ground phenomena (in the present case, phenomena associated with land use). However, sometimes

this is inadequate because of the density of the canopies. In almost any forest, however dense, there are areas where a LIDAR beam, at much shorter (optical) wavelengths, can penetrate to the ground through openings in the canopy that are of order of the wavelength, or greater, in size. An aircraft flying low and slow may receive a useful, although attenuated, return.

Any "Open Skies" program clearly depends on over flight permission negotiated by treaty. The data collected on any over flight must be made available to any participating nation, at a cost not to exceed the cost of reproduction. Smaller countries, unable to afford or field aircraft all over the world, will therefore get much more data on carbon worldwide than it could gather by itself. Whether such a program might be valuable depends strongly the level of access granted by treaty. This is beyond the scope of this study, but we point out the potential value and opportunity that "Open Skies" might present.

B.3 Future Directions

Monitoring of carbon fluxes solely from ground-based sensors will miss the crucial effects of upward convection, and therefore estimation of fluxes from ground concentrations will be hampered by meteorological uncertainty. A coordinated network of aircraft measurements on appropriate space and time scales may prove to have critical importance to models seeking to infer fluxes from concentrations. Current monitoring programs such as MOZAIC, with sensors mounted on scheduled commercial aircraft, are useful for vertical profile measurements near cities, where important carbon emitters are located. However, such planes fly at more or less constant altitudes in other regions, and other aircraft flying different routes and taking vertical profiles could be important.

Dramatic expansion of MOZAIC, even beyond IAGOS, may be extremely informative to models. We do not anticipate that "category-2" programs involving research flights will have a substantial production utility for de-

termining fluxes from models, although as R&D efforts for learning how to couple observations to fluxes via models they will be extremely valuable. We offer the vision of an "Open Skies" program for greenhouse gasses as a component of GHG treaties. Although less cost effective per kilometer than equipping commercial or military flights, such flights could observe CO_2 concentrations and infer fluxes directly, in a way to ascertain fluxes with minimal uncertainty and ambiguity.

Appendix B – References

[1] Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO_2, Science, **316** (5832), 17321735, doi:10.1126/science.1137004, 2007.

[2] Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rodenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe Part 1: mapping the atmospheric CO2 signals, Atmos. Chem. Phys. **7**, 34613479 (2007). http://www.atmos-chem-phys.net/7/3461/2007/.

[3] Xueref-Remy, I., Bousquet, P., Carouge, C., Rivier, L., Viovy, N., and Ciais, P.: Variability and 20 budget of CO2 in Europe: analysis of the CAATER airborne campaigns Part 2: Comparison of CO2 vertical variability and fluxes from observations and a modeling framework, Atmos. Chem. Phys. Discuss., bf 10, 42714304 (2010). http://www.atmos-chem-phys-discuss.net/10/4271/2010/.

[4] Marenco, Air and Space Europe **2**, no 3., p. 45 (2000).

[5] See the website `http://www.caribic-atmospheric.com/`.

C APPENDIX: Ocean Issues

Just as some countries, including the U.S., have suggested that a greenhouse gas treaty should give credit for uptake by forests and agriculture, so are some coastal states are likely to claim credit for uptake in their exclusive economic zones (EEZs), extending 200 nautical miles off shore. This raises a host of scientific and legal issues, not least of which is that outgassing nearly balances uptake, averaged globally (Figure 38). Because either can dominate locally, depending on ocean circulation, biological activity, and temperature, some waters will only add to a country's emissions if included in a treaty.

Technological changes are rapidly improving our ability to measure air-sea carbon fluxes and to monitor the ocean's carbon content. Both will be important if a carbon treaty has provisions for oceanic fluxes. Even if air-sea fluxes are not considered in a treaty, measurements of atmospheric CO_2 concentrations can be made close to most coastal states using ships of opportunity. Finally, use of Climate Process Teams by oceanographers offers a template for atmospheric scientists to break down barriers between modelers and measurers that appears to be limiting progress.

C.1 Air-sea CO_2 Fluxes

Air-sea CO_2 fluxes can be expressed simply as the product of two terms (C-11), the gas transfer velocity, k, and the difference in partial pressures in air, pCO_2^{a}, and water, pCO_2^{w}, across the air-water interface.

$$F = kK_0(pCO_2^{\mathrm{w}} - pCO_2^{\mathrm{a}}) \tag{C-11}$$

where F has units of mol m^{-2} year^{-1}, k has units of m s^{-1}, and K_0 is the aqueous-phase solubility (mol m^{-3} Pa^{-1}).

A legion of devils, however, lurk in the details, shown schematically in Figure 39. k varies with kinetic forcing of the air and sea boundary layers controlling dynamics of the gas transport. Kinetic forcing, in turn, is influenced

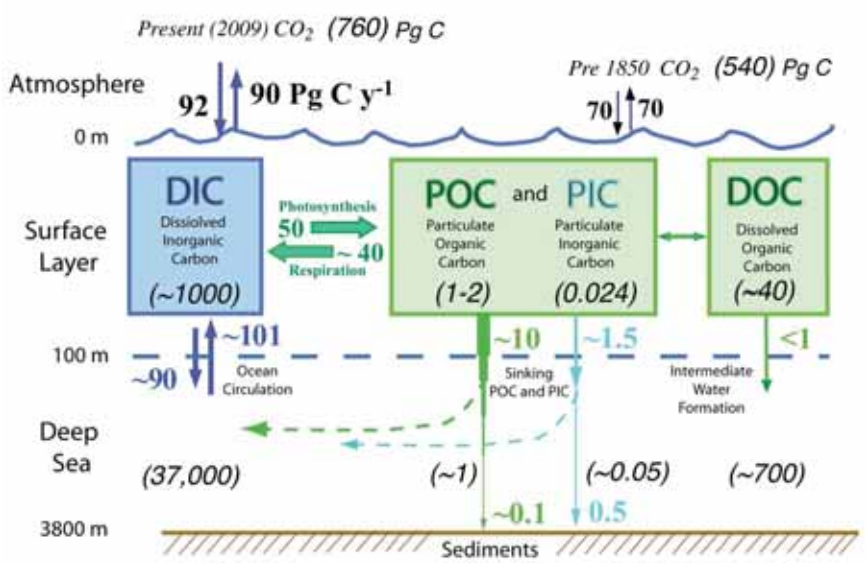


Figure 38: Schematic of the ocean's carbon inventory and fluxes [1]. Before 1850 air-sea CO_2 fluxes were in balance at 70 Pg C yr^{-1}. As a result of increasing atmospheric concentrations since 1850, last year, the ocean absorbed 2 Pg C yr^{-1} more than it gave up. Dissolved inorganic carbon (DIC) is the dominant form of carbon in the ocean. Only when POC and PIC are incorporated into the sediments is carbon removed from the ocean-atmosphere system.

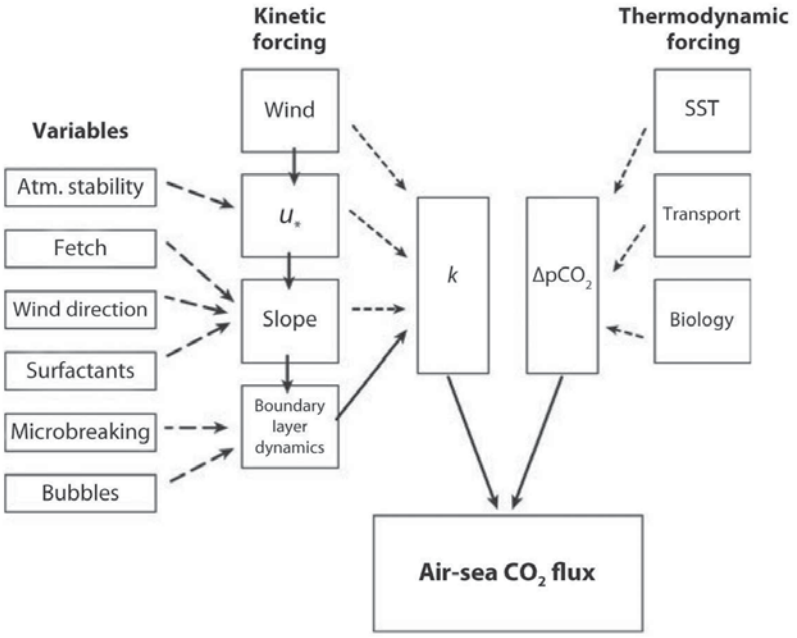


Figure 39: Simplified schematic of factors controlling air-sea CO_2 fluxes [6].

by secondary variables, such as atmospheric stability, direction and fetch of the wind, and the nature of breaking waves and the bubbles they create. Somewhat simpler, the partial pressure contrast depends on thermodynamic variables, such as sea surface temperature (SST), flushing of the interface by currents, and biology and its products, e.g., surfactant films on the surface.

Owing to the importance of these fluxes, intensive and successful field programs have been conducted, as carefully reviewed by [6]. Most importantly, fluxes were observed simultaneously from research ships applying covariance and profile methods, yielding consistent results. Covariance methods obtain correlations between gas concentrations, c, and vertical velocity, w,

$$F =< c'w' > +cw \tag{C-12}$$

where primes indicate fluctuations about averages, and w is the average velocity of dry air required to maintain its density in the presence of latent and sensible heat fluxes. The strong dependence of fluxes on wind speed led to fitting k as the sum of a constant and the cube of wind speed,

$$k_{660} = 3.3 + 0.026\ U_{10}^3 \quad \text{GasEx} - 1998 \tag{C-13}$$

$$= 8.2 + 0.014\ U_{10}^3 \quad \text{GasEx} - 2001 \tag{C-14}$$

where k_{660} is used for CO_2 because the Schmidt number (the ratio of diffusivity to kinematic viscosity, D/ν) is 660 for CO_2 at 20^{o}C in fresh water. U_{10} is the wind speed 10 m above the sea surface. GasEx-1998 was an observational program in the north Atlantic, where winds were moderately strong, and the cold ocean absorbed CO_2. GasEx-2001 went to the eastern Pacific, where winds were light and up-welling water near the equator was outgassing. In spite of being obtained in two very different oceanic regimes, the two fits are surprisingly consistent (Figure 40). The uncertainty of the covariance measurements, 2 mol m^{-2} year^{-1}, exceeds the average annual flux over the ocean of 0.5 mol m^{-2} year^{-1}, precluding its use over large areas of the ocean.

The profile method is based on measuring small concentration differences several meters apart vertically within the marine boundary layer when surface

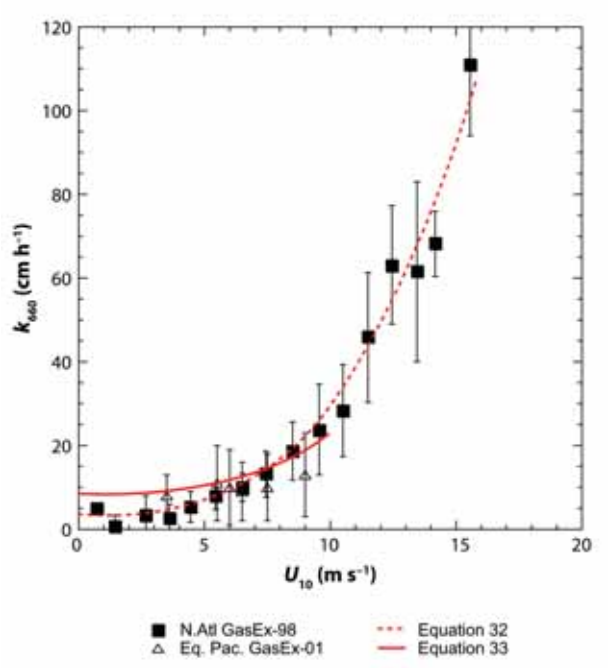


Figure 40: Observations and fits to CO_2 covariance fluxes from the north Atlantic (solid squares) and equatorial Pacific (open triangles), where winds were weaker than in the north Atlantic [6]. Equations 32 and 33 in the paper are reproduced above as (3) and (4).

momentum and buoyancy fluxes control dynamics of the marine boundary layer, allowing application of Monin-Obukhov similarity scaling. Then,

$$F = u_*^{\mathrm{a}}(c(z_1) - c(z_2))/R_{1,2} \tag{C-15}$$

where u_*^{a} is the friction velocity in air, and, during neutral conditions, the transfer resistance is $R_{1.2} = \kappa^{-1}\ln(z_2/z_1)$ with $\kappa = 0.41$ for von Karman's constant. During GasEx-1998 and GaxEx-2001 profile fluxes agreed with those from the covariance method.

C.2 Measuring CO_2 Concentrations Autonomously

Until recently, concentrations of carbon in the ocean could be measured only by researchers catching water samples from ships. This began to change several years ago when the National Ocean Partnership Program (NOPP) began

developing affordable bio-carbon sensors that can be added to autonomous ocean vehicles, followed by DOE support for the Carbon Flux Explorer which since 2001 has been obtaining autonomous measurements of particulate organic carbon (POC) and particulate inorganic carbon (PIC) [1]. Attached to the Sounding Oceanographic Lagrangian Observer (SOLO), the Explorer obtains profiles of the upper kilometer at a rate that can vary from a few hours to several weeks as the float drifts through the ocean, transmitting the previous profile every time it surfaces [1, 2]. These observations are concentrating on the rate at which POC and PIC fall out sink through the water column to ultimately settle on the sediments.

To obtain dissolved CO_2, it is necessary to measure the entire carbon system, requiring simultaneous measurement of nitrate and pH in addition to particulate carbon. Nitrate sensors have recently been deployed on profiling Lagrangian floats [3], and pH sensors with the required precision are being developed by the Monterey Bay Aquarium Research Institute (MBARI) and the University of Washington (S. Riser, personal communication, 2010). If the development succeeds, autonomous CO_2 measurements should be possible in a few years.

Once the sensors are developed, the issue for a carbon emissions treaty is whether and how they should be deployed. Global coverage is provided by Argo floats; somewhat more than 3,000 yield very good coverage of most of the ocean (Figure 41). Floats drift at depths of 2,000 m for about 10 days, when they take a profile while rising and remain on the surface while transmitting the data via Iridium. Because Argo floats cost about $15,000 apiece, adding carbon sensors would greatly increase the cost of a global program, but it may be possible to obtain adequate coverage by adding carbon sensors to a subset, say one in ten. Using global coupled atmosphere-ocean models to simulate different populations of carbon floats should guide the decision.

Drifting with ocean currents, floats are not constrained by Extended Economic Zones, but some countries, such as China, remain nervous about them

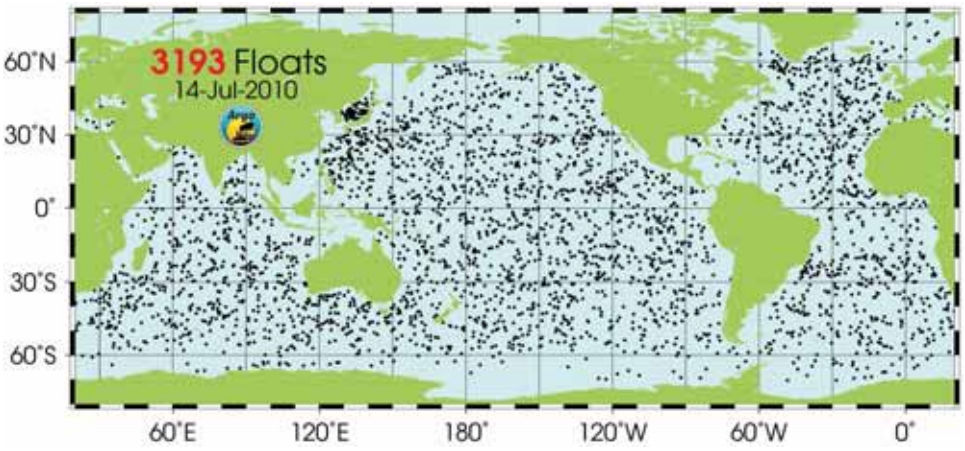


Figure 41: Distribution of Argo floats on 14 July 2010.

and would react if floats were deliberately released nearby to monitor their emissions, as the Naval Oceanographic Office learned after releasing some in the South China Sea. If negotiated as part of a treaty, however, gliders would provide a means of maintaining sensors at specified locations. Unpowered like floats, gliders can follow specified tracks or maintain station by navigating with tail fins during dives and climbs.

C.3 Monitoring Land Emissions from Ships

Seaborne sensors offer the prospect of monitoring emissions from coastal states, particularly uncooperative ones lacking nearby cooperative countries (Figure 42). [5] demonstrate how effectively data from a few offshore measurement stations used with atmospheric transport models can pinpoint sites emitting HCFC-22 and HFC-23 in China. Sensors on merchant ships regularly sailing near countries of concern may offer equally good or better possibilities for monitoring a carbon treaty as well as the Montreal Convention. At present, however, whether these measurements can be legally made is uncertain and contentious.

Although the U.S. has not ratified the Law of the Sea Treaty, our State Department requires U.S. oceanographic research ships to follow its provisions

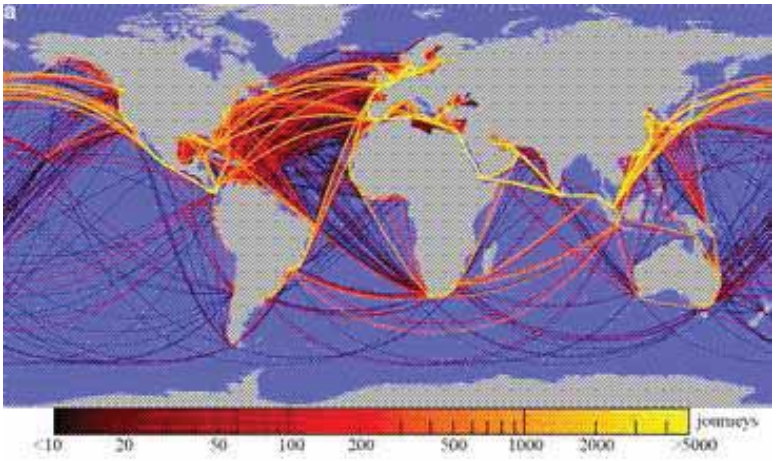


Figure 42: Shipping lanes colored by the number of journeys per year.

and request permission before making water measurements within Exclusive Economic Zones (EEZs) of any country. EEZs extend 200 nautical miles beyond the country's territorial waters, generally considered as extending 12 nautical miles from their coast. In some cases, countries consider their EEZ as beginning at the edge of their continental shelf, defined as the 200 meter isobath.

When diplomatic approval has not been obtained, NOAA research ships also stop atmospheric sensors, except standard meteorological balloons. Checking with some Law of the Sea experts, however, reveals that there is no basis under the treaty for limiting atmospheric research measurements in the EEZ (Prof. Ed Miles, Univ. Washington, personal communication), although grounds exist for banning them in territorial waters. Most pertinent to monitoring carbon, however, measurements in the EEZ could not be used in a legal case prosecuting a country for non-compliance. For measurements like these to have legal force, they must be established as part of a carbon treaty.

C.4 Process Teams

In some cases presented to us, poor communication between atmospheric modelers and measurers limits progress. A similar situation was recognized

several years ago by oceanographers concerned with climate models. Because the ocean is the primary heat sink of earth's climate system, correct representation in models of how mixing distributes heat vertically is essential for predicting climates beyond a few years ahead. Great progress has been made by observationalists and modelers, neither of whom were reading the other's literature or understanding their limitations and constraint.

To remedy the situation, an ad hoc group recommended forming mixing process teams, combining measurers and modelers, working together for several years to develop parameterizations of observational results considered ripe for incorporation into models. NSF's Physical Oceanography Program adopted the suggestion, requesting proposals for suitable topics and funding the first groups in 2003. Termed Climate Process Teams (CPTs), groups can recommend observations or numerical analysis needed to advance their topic in addition to developing parameterizations. Teams meet for two to three days at intervals of six months to a year, hiring a post doc or two and some doing limited data analysis or model runs. The object is to take topics ready for rapid progress once communication barriers are broken down, not to start new areas of research. Annual reports by the PI allow program managers to determine that progress is being made or decide that the topic or investigators are not ready and terminate funding.

One of the most successful, the Gravity Current Entrainment Climate Process Team (CPT) was established to find means of representing mixing in overflows ocean basin to another [4]. For instance, Mediterranean water overflowing the Strait of Gibraltar has a density that would take it to the bottom of the north Atlantic, but it ends up near 1,200 m depth because it mixes intensely while flowing across the continental shelf and slope. Neither the strait nor the overflow occupy more than 1 or 2 grid cells in coupled global models, making impossible correct representation of the mixing. Owing to the CPT, the best available information has been incorporated into useful parameterizations [4], and the need for additional measurements has been highlighted. A another round of proposals, addressing different topics

was recently funded.

Jason recommends that DOE develop a similar program bringing together modelers and observationalists on carbon-related topics.

Appendix C – References

[1] Bishop, J. (2010), Autonomous observations of the ocean's biological carbon pump, *Oceanography*, *22*(2), 183–207.

[2] Bishop, J., R. Davis, and J. Sherman (2002), Robotic observations of dust storm enhancement of carbon biomass in the north Pacific, *Science*, *298*, doi:10.1126/science.1074,961.

[3] Kohnson, K., S. Riser, and D. Karl (2010), Nitrate supply from deep to near surface waters of the north Pacific subtropical gyre, *Nature*, *in press*.

[4] Legg, S., et al. (2009), Improving oceanic overflow representation in climate models, *Bull. Am. Met. Soc.*, *90*(5), 657–670.

[5] Stohl, A., et al. (2010), Hydrochlorofluorocarbon and hydrofluorocarbon emissions in east asia determined by inverse modeling, *Atmos. Chem. Phys.*, *10*, 3545–3560.

[6] Wanninhkof, r., W. Asher, D. Ho, C. Sweeny, and W. McGillis (2009), Advances in quantifying air-sea gas exchange and environmental forcing, *Annu. Rev. Mar. Sci.*, *1*, 213–244.

D APPENDIX: Isotopic Signatures

D.1 Dilution by fossil carbon

Fossil carbon contains no ^{14}C and is about 18 parts per thousand (‰) depleted in ^{13}C compared to atmospheric carbon. An elementary calculation shows that adding a mass δM_f of fossil carbon to an atmosphere containing a mass M_a of pre-industrial carbon of isotopic fraction x_n and M_f of fossil carbon of isotopic fraction x_f produces, to first order in $\delta M_f/(M_a, M_f) \ll 1$, a mixture of isotopic fraction

$$\frac{\delta x}{x} \approx \delta M_f (x_f - x_n) \frac{M_a}{(M_a + M_f)(M_a x_n + M_f x_f)}. \tag{D-16}$$

For ^{14}C, $x_f = 0$ and this reduces to the obvious result

$$\delta_{14} \approx -\frac{\delta M_f}{M_a + M_f}. \tag{D-17}$$

For ^{13}C, $x_f - x_n \approx -0.018 x_n$, so that to lowest order in this small quantity

$$\delta_{13} \approx -0.018 x_n \delta M_f \frac{M_a}{(M_a + M_f)^2}. \tag{D-18}$$

At present $M_a/(M_a + M_f) \approx 0.7$, so that $\delta_{13} \approx 0.013 \delta_{14}$, and is not likely to be measurable.

D.2 Carbon 14

Carbon dioxide is a common constituent of the atmosphere that is exchanged with the soil and oceans by natural processes modulated diurnally, seasonally, and annually. It is difficult to measure anthropogenic increases on short time scales, especially when the emissions must be spatially resolved. The input of CO_2 by fossil-fuel burning, which dominates the anthropogenic emissions, bears distinctive signatures in the distribution of carbon and oxygen isotopes and in combustion products other than CO_2 itself.

The gold standard in this respect is ^{14}C, or rather $^{14}CO_2$, because its radioactive decay has removed it almost completely from fossil fuels, whereas it is produced in the atmosphere and biosphere by cosmic-rays from which underground reservoirs of fossil fuels are naturally shielded. On the other hand, the 5730-year half-life of ^{14}C is sufficiently long that its abundance in most living things is in equilibrium with that in the atmosphere, apart from some isotopic fractionation that accompanies photosynthesis. Thus, the burning of fossil fuels dilutes the mixing ratio of ^{14}C relative to the stable isotopes ^{12}C and ^{13}C in the air, in plant and animal tissue, and in agricultural products. This dilution, called the Suess effect, is conventionally quantified by $\Delta^{14}C$, defined in such a way [1] as to correct for photosynthetic fractionation and usually expressed in parts per thousand (‰) such that $(\Delta^{14}C)_{\rm fossilfuels} = -1000‰$ (Figure 43). If $[CO_2]_{\rm m}$ is the total measured molar fraction of CO_2 in an air sample summed over isotopic species, then the molar fraction due to fossil-fuel burning is

$$[CO_2]_{\rm ff} = [CO_2]_{\rm m} \ \times \ \frac{\Delta^{14}C_{\rm bkg} - \Delta^{14}C_{\rm m}}{\Delta^{14}C_{\rm bkg} + 1000‰}\,, \qquad \text{(D-19)}$$

in which "bkg" refers to a background level that need not be pristine.

The long-term trend in the dilution of ^{14}C exhibited by Figure 43 is not due solely to fossil-fuel emissions. Atmospheric nuclear testing nearly doubled the atmospheric inventory of ^{14}C, an event from which $\Delta^{14}C$ is only now recovering (Figure 44). The rate of recovery is obviously much too rapid to be explained by radioactive decay. Thus, the nuclear experiments of the Cold War have been used to calibrate the rate of exchange of CO_2 between the atmosphere and the land and sea [7].

Plants, including agricultural crops, have been used as natural sample collectors for measuring $\Delta^{14}C$ and hence $[CO_2]_{\rm ff}$. An example is shown in Figure 45, taken from a study reported in [8]. The authors collected samples of wild winter annual grasses from 128 sites across California. These were analyzed for their $\Delta^{14}C$ content, with the results shown in the left hand panel of Figure 45. The numbers on the map are hard to make out in this repro-

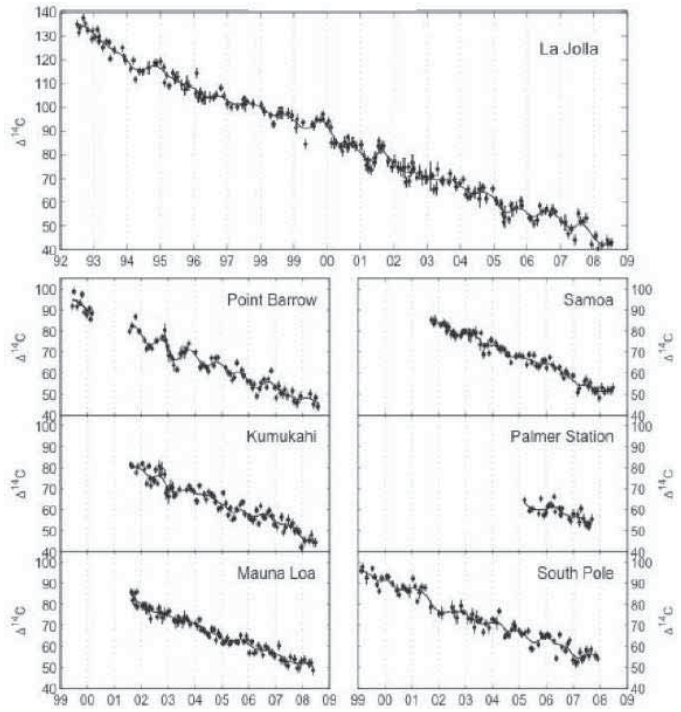


Figure 43: Δ^{14}C time series from the seven global sampling stations of the Scripps Institution of Oceanography [2, 3].

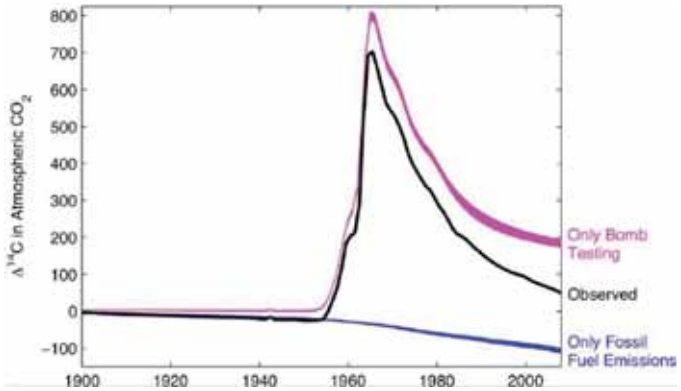


Figure 44: Production of ^{14}C by atmospheric nuclear testing [3].

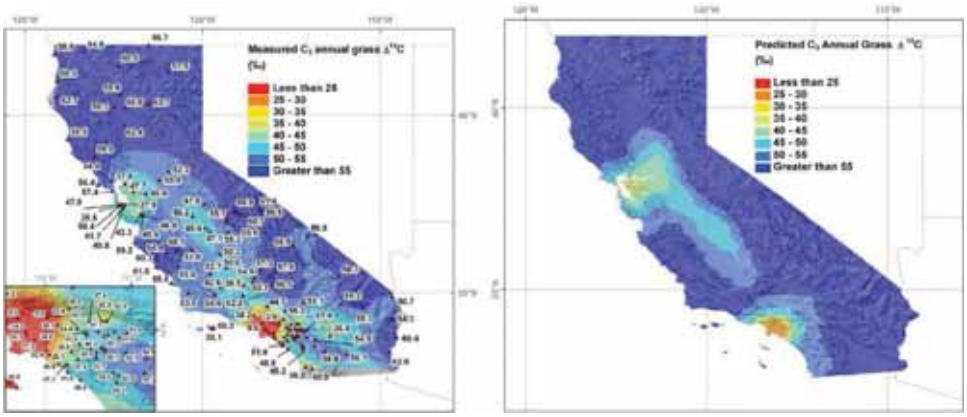


Figure 45: Measured (*left*) and predicted (*right*) Δ^{14}C in California. The measurements were based on samples of wild grasses, and the models used (measured) NO_x concentrations as proxies for the CO_2 sources [8].

duction, but the colors and accompanying legend show that $\Delta^{14}C < 25‰$ in Los Angeles (*red*), as compared to a background level $\Delta^{14}C \approx 60 \pm 2.5‰$ on the North Coast and comparably remote areas. (The background level is positive—higher than the natural equilibrium between cosmic-ray production and ^{14}C decay—because of the residual influence of atmospheric nuclear tests.) The predictions shown in the right hand panel were made with models for atmospheric transport and grass growth; the transport model required specification of the fossil-fuel emissions, for which previous measurements of NO_x on a 36-km grid, scaled to a bottom-up inventory of statewide fossil-fuel emissions, were used as proxies. In this study at least, the level of agreement between model and measurement is quite impressive:

> Average measured Δ_g [Δ^{14}C of annual grasses] levels in Los Angeles, San Francisco, the Central Valley, and the North Coast [of California] were 27.7±20.0, 44.0±10.9, 48.7±1.9, and 59.9±2.5‰, respectively. Model predictions reproduced regional patterns reasonably well, with estimates of 27.6 ± 2.4, 39.4 ± 3.9, 46.8 ± 3.0, and 59.3±0.2‰ for these same regions and corresponding to fossil fuel CO_2 mixing ratios (C_f) of 13.7, 6.1, 4.8, and 0.3 ppm [8].

Note that in Eq. 1-1 $\Delta^{14}C_{bkg} \approx 60 \pm 2.5‰$ so that the measured North Coast

$\Delta^{14}C_m = 59.3 \pm 0.2‰$ corresponds to negligible fossil fuel sources, and that regions with significant sources of fossil carbon have smaller $\Delta^{14}C$.

This shows what can be achieved in an economically well-developed, fully accessible region by a judicious combination of dense sampling and careful modeling. Similar studies have been carried out across the U.S. using maize [9], and in Europe using wine [10]. The quoted uncertainties in the inferred concentrations of CO_2 are on the order of 1 ppm in all three studies.

D.3 Relevance to Treaty Verification

The important question for the present study is, to what accuracy could the fossil-fuel CO_2 fluxes or concentration be determined via $\Delta^{14}C$ in a "non-cooperative" country? There is, first of all, the problem of obtaining the necessary samples. Possibly this could be done by relying on agricultural imports, including not only food products but also lumber, bamboo, cotton, *etc.* It would be necessary to know where the original plants grew, to an accuracy determined by the requirements of the model: perhaps a few tens of kilometers. This information may be difficult to obtain for most products.

For some luxury goods, however, the specific growing region is proudly advertised. Such is the case with French and German wines (although the grapes in any given wine are often drawn from several vineyards), but Europe must be counted as a cooperative region. It is also the case for some high-end Indian teas; unfortunately, the main tea-producing regions of that country are concentrated in the extreme north, east, and south, rather than being spread throughout it. Still, if some proxy analogous to the NO_x used in Ref. [8] in California can be employed to estimate the spatial distribution of sources (e.g., CO, sulfur, or particulate emissions), then in combination with transport models and appropriate meteorological data, accurate $\Delta^{14}C$ measurements from a few such regions might allow the year-to-year trend in net fossil-fuel CO_2 emissions to be usefully constrained. The extent to which this is realistic could be clarified by re-analyzing the studies in Refs. [8, 9],

and if necessary by performing more such studies within cooperative regions.

The principal fossil-fuel emitters other than the U.S.—India and China—though perhaps not fully cooperative with all the desires of a carbon-treaty verifier, are far from being closed societies. Foreigners can travel within these countries rather freely, and so it might not be very easy for their governments to prevent the direct collection of plant samples from rural areas without disrupting the international commerce and tourism upon which these countries rely for much of their economic growth. Meteorological data are restricted mainly by the prevalence and sophistication of measurement and communications infrastructure rather than by secrecy.

The utility of Δ^{14}C observations from *sparse* but fully controlled stations to monitor emissions reductions in Europe has been addressed [11]. Analyzing two decades of data collected at two sites in Germany, they observed year-to-year fluctuations in the inferred (from local Δ^{14}C compared to tropospheric background) fossil-fuel CO_2 concentration that cannot plausibly be due to fluctuations in the actual emissions (because these are believed to be rather constant), and which they therefore attribute, in roughly equal measure, to a combination of measurement uncertainties and biogenic processes on the one hand, and to unmodeled transport effects on the other. On this basis, they estimate that emissions reductions in southern Germany and eastern France of $\geq 26\%$ over a five year period could be detected with 95% confidence from their more remote site (Schauinsland, in the Black Forest), and $\geq 7\%$ from the more polluted site (Heidelberg). Similarly, [12] estimated that net annual global $[CO_2]_{\mathrm{ff}}$ emissions, as traced by observations of Δ^{14}C by a network of 8 remote sites spread across both hemispheres, are uncertain at the 25% level; they based this conclusion on the extent to which models of these data are able to explain the north-south difference in Δ^{14}C, and particularly its inter annual variation.

Biases in the inference of local $[CO_2]_{\mathrm{ff}}$ concentration (rather than flux) from Δ^{14}C measurements associated with the choice of background site, where the latter is supposed to be sufficiently remote as to represent a hemispherical

average were examined in Ref.˜[]t09. These authors concluded that it is possible to identify and correct for these biases, and that a single $\Delta^{14}C$ measurement of 2‰ precision (present state of the art) would be sufficient to infer $[CO_2]_{ff}$ concentration to 1 ppm. They freely acknowledge, however, that the inference of $[CO_2]_{ff}$ *fluxes* from concentrations is dominated by modeling uncertainties.

D.4 The Suess Effect

In 1955 Suess discovered that the ^{14}C concentration of modern carbon (atmospheric and photosynthetic) is less than that of pre-industrial carbon, allowing for radioactive decay. He explained this as a consequence of the dilution of atmospheric CO_2 by fossil carbon.

The atmospheric carbon cycle is more complicated than simple addition of anthropogenic carbon to a pre-industrial level of CO_2. Carbon is exchanged among the atmosphere, biosphere, and ocean, and these exchanges are not understood quantitatively. The importance of these processes is shown by the results of [14] (among many other studies). These authors found $\delta_{14} = -20‰$ in comparing pre-industrial levels with levels from 1950 (after 1950 large positive δ_{14} are found because of production by atmospheric nuclear explosions). Yet modern estimates of pre-industrial (275 ppm) and 1950 (310 ppm) CO_2 levels would imply, if atmospheric dilution were the only relevant process, $\delta_{14} \approx -113‰$.

In order to understand the quantitative relation between anthropogenic emissions and atmospheric CO_2 levels, it is necessary to model these exchanges. Their influence is also evident in the decline of atmospheric ^{14}C since the cessation of atmospheric nuclear explosions at a rate much greater than would be produced by anthropogenic dilution and radioactive decay alone (see Figs. **??** and 44).

If there is not sufficient time for significant exchange with the ocean or bio-

sphere these processes will not affect the downwind signature of regional anthropogenic emissions. Although that might be expected to be the case for sensors only hundreds of km downwind that sense emissions within the previous day, this may not always be the case. For example, a sensor on a buoy will be immersed in the marine micro-boundary layer in which air-sea exchange is rapid because of its thinness; it may observe δ_{14} and CO_2 concentrations significantly different than the average of the 1–2 km atmospheric boundary layer. In addition, scavenging by rain may be a significant means of removing CO_2 from the atmosphere, and may affect its isotopic abundances even near its source.

D.5 Measuring ^{14}C

A difficulty with ^{14}C as a tracer of fossil-fuel carbon is the rarity of this isotope, approximately one part in a trillion compared to ^{12}C. Remote sensing of ^{14}C (*e.g.*, by absorption spectroscopy) is impractical. Instead, samples must be collected and taken to a laboratory where sensitive equipment measures the number of ^{14}C atoms in the sample by counting radioactive decays. At its natural concentration (even after molecules other than CO_2 are removed from the air sample by chemical processing), the count rate in samples of practical size is low enough that days to weeks are required for an accurate measurement of $\Delta^{14}C$. Alternatively, accelerator mass spectrometry (AMS) counts ^{14}C atoms by separating them from the other isotopes without relying on radioactive decays; analysis times per sample are minutes to hours, and the samples are only milligrams of carbon as opposed to grams for direct counting sample. Accuracies at the level of a few parts per thousand are typically reported, and throughput is several thousand samples per year, ranging up to more than ten thousand for the LLNL AMS [5, 6]. AMS equipment is bulky and power-hungry. Compact devices that are several meters on a side are being developed, including gas-source systems [17].

D.6 Stable Carbon Isotopes

The $^{14}C/^{12}C$ ratio is the traditional means of distinguishing CO_2 produced by combustion of fossil *vs.* modern carbonaceous material. However, an additional and independent datum is valuable. Consider, for example, a treaty the commits a country to limit its combustion of fossil fuels in electric generation and transportation. In verifying compliance it would be necessary to distinguish those sources of ^{14}C-free CO_2 from smoldering coal or old peat beds (or old soil carbon). It would also be necessary to distinguish between rapid combustion in modern generators and vehicles and slower combustion of coal in traditional fireplaces. These distinctions might require the evidence from oxygen isotopic abundances, because all these fuels are free of ^{14}C.

Isotopes fractionate at sources and sinks, through phase changes, and in chemical reactions as a result of kinetic and thermodynamic effects at the molecular level. Variations in the bulk isotopic composition of CO_2 (*i.e.*, $^{13}C/^{12}C$, $^{18}O/^{16}O$, $^{17}O/^{16}O$) play a potentially important role in quantifying carbon sources and sinks. Fractionation of these elements during chemical, physical, and biological processes results in potentially unique labels. ^{13}C, with roughly 1.1% the abundance of ^{12}C, has been useful for constraining the evolution of the chemical composition of the atmosphere and assessing and quantifying the uptake of CO_2 in the oceans, land, and biosphere, as well as distinguishing natural from anthropogenic carbon in the atmosphere. The primary issue is that the individual stable isotopes of carbon and oxygen from different sources in general have overlapping compositions and be affected by subsequent exchange reactions. However, the overlap in isotopic compositions from different sources typically is not identical, and additional constraints on CO_2 sources can be obtained by taking all the stable isotope data into consideration and coupling the results with concentration measurements and different flux models.

The isotopic composition of a sample is conventionally defined as $\delta = (R_{sample}/R_{standard} - 1) \times 1000$. For the isotopes of interest here, the standards are the

Vienna Pee-Dee Belemnite (V-PDB) carbonate for carbon isotopes and the Vienna Standard Mean Ocean Water (V-SMOW) for oxygen (and hydrogen; see Ref. [18]). Representative values for δ^{13}C include -24 ‰ and -27 ‰ for coal and lignite, respectively; -20 ‰ to -75 ‰ (mean of -41 ‰) for natural gas; approximately -5 ‰ for volcanic CO_2, which is derived from the mantle; a spread in δ^{13}C values centered around -25 ‰ for crude oils, which is derived from the precursor molecules [18].

Because biologically derived carbon is light (negative δ^{13}C); the change in δ^{13}C of atmospheric CO_2 of $\sim$ 1.5‰ over the past 100 years is consistent with the rise of anthropogenic emissions. However, there are large seasonal and diurnal variations, which are reflected in the isotopic "Keeling curve" shown in Figure 46. The seasonal cycle dominates the terrestrial biosphere of the northern hemisphere, and the long term trend and difference between the hemispheres in these datasets reflect the effects of fossil fuel emission.

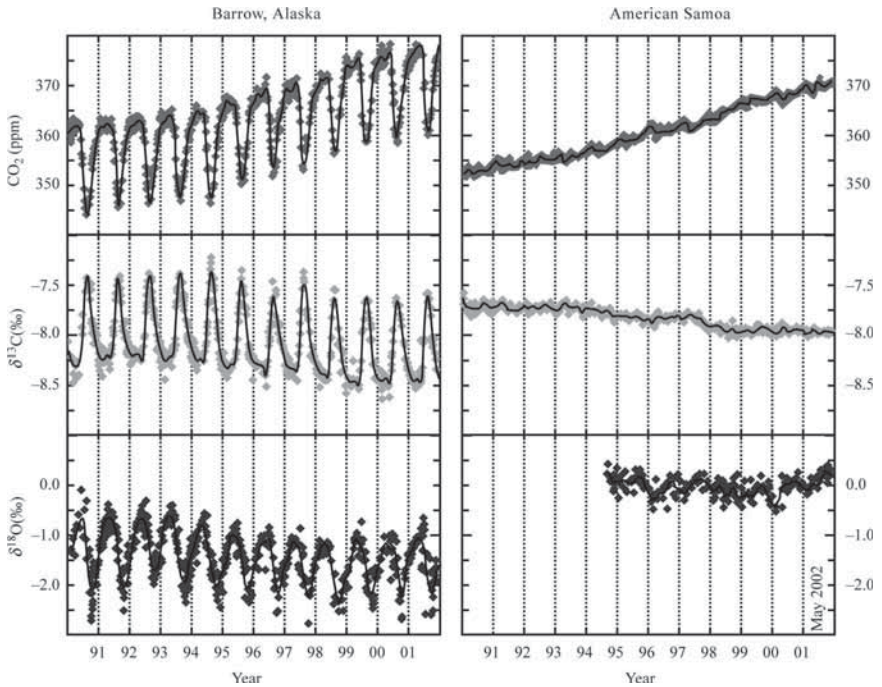


Figure 46: Seasonal and long-term trends in concentrations of δ^{13}C and δ^{18}O of atmospheric CO_2 measured in the northern and southern hemispheres [18].

This rise in atmospheric anthropogenic CO_2 has infiltrated the isotopic composition of other pools. The ocean-atmospheric isotopic fluxes vary significantly both on spatial and temporal scales (Figure 47). In contrast to the effect of biological processes, atmospheric CO_2 depletes surface ocean dissolved inorganic carbon in ^{13}C. The exchange of that isotope with the ocean is based predominantly on physical equilibrium, which gives rise to two exchange fluxes that are isotopically distinct with different contributions to the atmospheric ^{13}C. The ^{13}C values of fossil fuels and the land-use produced CO_2 in the atmosphere are imprinted with the photosynthetically produced carbon, which discriminates against ^{13}C. There is no evidence for fractionation of carbon isotopes on burning of fossil fuels, so the isotopic distribution from the sources is assumed to be that imparted to the atmosphere from the fuels. Superimposed on the dominant source-sink fractionation are disequilibrium effects which arise from different time scales over which isotopic equilibrium is reached for the ocean-atmosphere system and mineralization of total ocean organic matter [18].

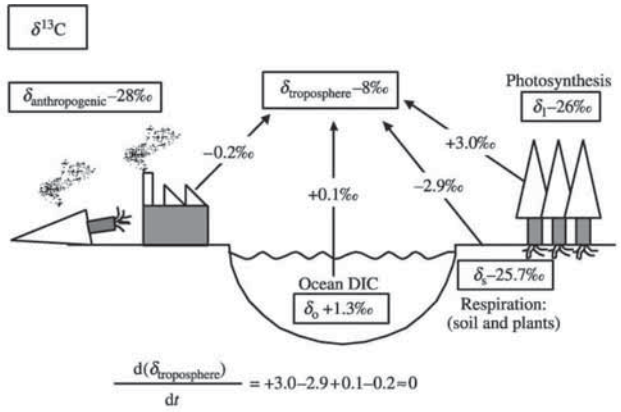


Figure 47: Annual mean forcings on $\delta^{13}C$ values of CO_2 [18].

D.7 Oxygen Isotopes

There are three stable isotopes of oxygen: ^{16}O, ^{17}O, and ^{18}O. $\delta^{18}O$ of CO_2 derived from volcanoes and permafrost depends on $\delta^{18}O$ of the water present in the system, *i.e.,* the $\delta^{18}O$ of the magmatic and soil water, respectively.

The $\delta^{18}O$ of CO_2 produced by combustion of fossil fuel depends on the C/H in the fuel and equilibration temperature between CO_2 and H_2O exhaust. A complicating factor is that the ^{18}O of CO_2 can be reset during its transport in air (through exposure to water, especially leaf water). There can be isotopic fractionation associated with photochemistry [*e.g.,* $O(^1D) + CO_2$], some of which can give rise to non-mass dependent isotope effects [18]. Aircraft measurements of the composition of stratospheric CO_2 with respect to all three oxygen isotopes in whole air samples show anomalous enrichments in ^{17}O and ^{18}O [19]. The correlation of the isotope anomaly with simultaneous N_2O measurements demonstrates that ^{17}O in CO_2 is a potentially useful tracer.

D.7.1 Oxygen Isotope Diagnostics

It has long been known [20] that oxygen chemistry produces measurable fractionation of its stable isotopes. Usually this fractionation may be approximated as proportional to the mass difference (or a power of its reciprocal), a natural consequence of the small isotopic shifts in vibrational frequencies and therefore in the reaction energies, but in some cases the fractionation is independent of the mass difference, or has a more complex dependence. Oxygen isotopic fractionation has been the subject of hundreds of publications and is discussed in a recent textbook [21].

These fractionations distinguish among various low temperature (respiratory) oxidation pathways, between them and high temperature (burning) oxidation, and among various photosynthetic pathways. There is an extensive research literature. In general, oxygen isotope ratios in the product CO_2 depend on both the temperature of reaction [22] and on the specific biochemical pathway. Studies include the differential isotopic fractionation between cyanide-resistant and cyanide-sensitive respiration [23], and the isotopic fractionation in combustion [24].

Perhaps surprisingly, isotopic fractionation in combustion is significant, con-

trary to the expectation that at high temperatures small isotopic differences in bond energies would not lead to fractionation. Even low temperature respiration requires large energies to break covalent bonds, and the fact that fractionation occurs in respiration is a reason to anticipate measurable isotopic effects in combustion. The *rate* of combustion may influence the isotopic abundances of the product CO_2 [24]. This offers, for example, the prospect of using these abundances to distinguish between rapid burning of fossil fuels and smoldering combustion of agricultural waste, soil humus, and coal beds.

Not only are there significant differences in the mass dependence of the oxygen isotope fractionation between various biochemical pathways, both in oxygen production (photosynthesis) and in oxygen consumption (respiration), but the differential fractionation of the three stable isotopes depends on the process [25]. This result implies that measurement of the two independent isotopic ratios ($^{17}O/^{16}O$ and $^{18}O/^{16}O$) provides two *independent* constraints on the production of the CO_2, rather than only one, as would be the case if all fractionations had the same dependence on isotopic mass ratio.

Oxygen isotope ratios in the products of photosynthesis may also depend on the oxidative pathway [26]. These contribute to the isotopic ratios in the products of oxidation, either biochemical or by combustion (some of the oxygen in the CO_2 is atmospheric, but some of it derives from oxidized compounds, such as cellulose, that contain oxygen).

It is therefore possible that, even if there is no immediate application to monitoring CO_2 emissions, study of oxygen isotopic abundances in CO_2 would contribute indirectly by aiding the understanding of global the carbon cycle.

D.8 Isotope Clumping

Although the overlap and variation of ratios of the above stable isotopes limit their utility as unique fingerprints for specific sources, their 'clumping'

in molecules can provide additional information. Clumping of isotopes is the thermodynamic stabilization of orderings of isotopes in molecules. This effect enhances the formation (relative to a statistical distribution) of specific isotopomers, which are molecules with the same number of isotopes but in different positions (Table 5) [27].

Table 5: Isotopologues

Mass	Isotopologue	Relative abundance
28	$^{14}N_2$	99.3%
29	$^{14}N^{15}N$	7.30×10^{-3}
30	$^{15}N_2$	1.34×10^{-5}
32	$^{16}O_2$	99.5%
33	$^{17}O^{16}O$	7.56×10^{-4}
34	$^{18}O^{16}O$	4.00×10^{-3}
	$^{17}O_2$	1.44×10^{-7}
35	$^{18}O^{17}O$	1.52×10^{-6}
36	$^{18}O_2$	4.00×10^{-6}
44	$^{12}C^{16}O_2$	98.4%
45	$^{13}C^{16}O_2$	1.11×10^{-2}
	$^{12}C^{17}O^{16}O$	7.48×10^{-4}
46	$^{12}C^{18}O^{16}O$	4.00×10^{-3}
	$^{13}C^{17}O^{16}O$	8.40×10^{-6}
	$^{12}C^{17}O_2$	1.42×10^{-7}
47	$^{13}C^{18}O^{16}O$	4.44×10^{-5}
	$^{12}C^{18}O^{17}O$	1.50×10^{-6}
	$^{13}C^{17}O_2$	1.60×10^{-9}
48	$^{12}C^{18}O_2$	3.96×10^{-6}
	$^{13}C^{18}O^{17}O$	1.68×10^{-8}
	$^{13}C^{18}O_2$	4.45×10^{-8}

Notes: Nominal cardinal mass in AMU; assuming $^{15}N/^{14}N$ ratio equal to atmospheric N_2, $^{17}O/^{16}O$ and $^{17}O/^{16}O$ ratios equal to the V-SMOW standard, and $^{13}C/^{12}C$ ratio equal to the V-PDB standard.

Early measurements of this effect in carbon dioxide suggests it could be relevant to treaty monitoring of anthropogenic emissions. The ratio $^{16}O^{13}C^{18}O/^{16}O^{12}C^{16}O$ ($^{47}CO_2/^{44}CO_2$) can distinguish CO_2 of automobile exhaust from background signals [34]. The temperature of formation was recorded in the $^{47}CO_2/^{44}CO_2$

ratio. Again, CO_2 exchanges oxygen readily with water, so the temperature recorded by the emitted gas was lower than the combustion temperature due to reactions with water in the tailpipe, which was $\sim 100^o$C. Nevertheless, the signal for CO_2 was distinct and could be readily measured using standard mass spectrometric techniques.

Subsequent measurements have shown an enrichment of $^{47}CO_2/^{44}CO_2$ in the stratosphere, perhaps associated with the photochemically promoted $O(^1D)$ + CO_2 reaction [28]. The measurements provide evidence for an exchange of O-isotopes between highly enriched ozone and CO_2 in the stratosphere. Despite the low concentrations, the isotopic signals are large and measurable, and show a $^{47}CO_2$ clumping signal. CO_2 derived from high temperature processes is relatively depleted in clumping effects, compared to CO_2 derived from low temperature processes; notably, the clumped isotopic composition of respiratory CO_2 is depleted relative to the equilibrium values [33].

As for the oxygen isotopes alone, clumped isotope signals are also affected by water reactions. However, the resetting of the clumped isotope (and oxygen isotope) compositions of atmospheric CO_2 though the exchange reaction with water does not occur instantaneously. Thus, the characteristic isotopic signals of different CO_2 sources masked by the exchange are not completely removed [29], thereby allowing inferences about carbon sources. Temporal monitoring of CO_2 isotope compositions (*e.g.,* observed seasonal and diurnal variations) have been able to measure these signals [29].

The clumped isotope composition of CO_2 could be used as a tool to distinguish CO_2 derived from high temperature processes (*e.g.,* volcanoes and power plants) and low temperature processes (*e.g.,* permafrost); the higher temperature, the smaller clumping effect. The clumping associated with other anthropogenic emissions *vs.* that of natural sources has not yet been investigated in detail. The technique opens up the prospects for direct measurements that could be diagnostic of specific combustion sites.

The mass resolving power of conventional gas source mass spectrometers

($M/\Delta M \approx 1000$) is sufficient to identify clumping signals in CO_2. Distinguishing among isotopologues of other greenhouse gases can require higher resolution. For methane, calculations of relative energies of $^{12}CD_2H_2$ and $^{12}CDH_3$ show that measurements of their ratio would permit the discrimination of formation temperatures of $\sim 200°C$ to be determined, thereby allowing different methane sources to be identified. This type of measurement requires $M/\Delta M = 60{,}000$, which is possible with high resolution tandem mass spectrometers being built [33]. Similar measurements could quantify sources of other greenhouse gases as well as higher hydrocarbons. It is useful to point out here that coupling of these measurements with fluxes up-welling and down-welling gas would provide additional constraints.

Appendix D – References

[1] Stuiver, M., and Polach, H., *Radiocarbon* **9**, 355–363 (1977).

[2] Guilderson, T. P., *et al.* LLNL-TR-411412 (2009).

[3] Guilderson, T. P., briefing to Jason, June 2010.

[4] Graven, H., Guilderson, T. P., and Keeling, R. F. *J. Geophys. Res.* submitted (2010).

[5] Roberts, M. L., *et al. Nucl. Instr. Meth. Physics Res. B* **123**, 57–61 (1997).

[6] Southon, J., *et al. Radiocarbon* **46**, 41–49 (2004).

[7] Naegler, T., and Levin, I. *J. Geophys. Res.* **114**, D17303 (2009).

[8] Riley, W. J., *et al. J. Geophys. Res.* **113**, G04002 (2008).

[9] Hsueh, D. Y., *et al. Geophys. Res. Lett.* **34**, L02816 (2007).

[10] Plastra, S. W. L., *et al. J. Geophys. Res.* **113**, D21305 (2008).

[11] Levin, I., and Rödenbeck, C., *Naturwissenschaften 95*, 203–208 (2008).

[12] Levin, I., *Tellus* **62B**, 26–46 (2010).

[13] Turnbull, J. C., *et al. J. Geophys. Res.* **114**, D22302 (2009).

[14] Stuiver, M., and Quay, P. D. *Earth & Planet. Sci. Lett.* **53**, 349–362 (1981).

[15] R. Hellborg and G. Skog, Accelerator mass spectrometry. *Mass Spectrom. Rev.* **27**, 398-427 (2008).

[16] S. W. Pacala et al., Verifying greenhouse gas emissions: methods to support international climate agreements (National Resource Council, Washington, DC, 2010); prepublication version.

[17] M. L. Roberts et al., Progress on a gas-accepting ion source for continuous-flow accelerator mass spectrometry. *Nucl. Instrum. Meth. B* **259**, 83-87 (2007).

[18] D. Yakir, The stable isotopic composition of atmospheric CO_2, *Treatise on Geochemistry*, Vol. 4, 175-212 (Elsevier Ltd., 2003).

[19] K. A. Boering et al., Observations of the anomalous oxygen isotopic composition of carbon dioxide in the lower stratosphere and the flux of the anomaly to the troposphere. *Geophys. Res. Lett.* **31** L03109, doi:10.1029/2003GL018451 (2004).

[20] Lane. G. A., and Pole, M. *Science* **123**, 574–576 (1956).

[21] Lambers, H., Chapin III, F. S. and Pons, T. L. *Plant Physiological Ecology* 2nd ed. (Springer 2008).

[22] Schleser, G. H. *Rad. & Env. Biophys.* **17**, 85–93 (1979).

[23] Guy, R. D., *et al. Planta* **177**, 483–491 (1979).

[24] Schumacher, M. *et al. Atmos. Chem. Phys. Discuss.* **8**, 18993–19034 (2008).

[25] Helman, Y. *et al. Plant Phys.* **138**, 2292–2298 (2005).

[26] Luo, Y.-H., and Sternberg, L. Da S. L. *J. Expt. Bot.* **43**, 47–50 (1992).

[27] J. M. Eiler, "Clumped-isotope" geochemistry - The study of naturally-occurring, multiply-substituted isotopologues. *Earth Planet. Sci. Lett.* **262**, 309-327 (2007).

[28] L. Y. Yeung et al., Large and unexpected enrichment in stratospheric $^{13}C/^{12}C$ and its meridional variation. *Proc. Nat. Acad. Sci.* **106**, 11496-11501 (2009).

[29] H. P. Affeck, X. Xu, and J. M. Eiler, Seasonal and diurnal variations of $^{13}C^{18}O^{16}O$ in air: initial observations from Pasadena, CA. *Geochim. Cosmochim. Acta* **71**, 5033-5043 (2007).

[30] Kalckar, F., and Teller, E. *Proc. Roy. Soc. Lond.* A **150**, 520–533 (1935).

[31] D. Archer, Methane hydrate stability and anthropogenic climate change. *Biogeosciences* **4**, 521-544 (2007).

[32] R. A. Houghton, Balancing the global carbon budget. *Ann. Rev. Earth Planet. Sci.* **35**, 313-347 (2007).

[33] W. Guo, private communication.

[34] H. P. Affeck and J. M. Eiler, Abundance of mass $^{47}CO_2$ in urban air, car exhaust, and human breath. *Geochim. Cosmochim. Acta* **71**, 1-12 (2006).

E APPENDIX: Statistical Data Assimilation

This appendix is a summary of methods for transferring information from observations of a physical (or other) experimental or field system to a model developed by thinking (separately) about the dynamics of the underlying system. In geophysical contexts this is known as **data assimilation**. The goal is the estimation, conditioned on the data, of the fixed parameters and unobserved state variables of the model at the end of an observation period $t = T$. This allows predictions of the model state $\mathbf{x}(t > T)$ using the model dynamics. It can also be used to test the model and suggest where to look in state space to improve the model.

The formulation of data assimilation given as an exact path integral over the model state through the observation period. The "saddle path" evaluation of the integral representation of the statistical questions is the familiar 4DVar formulation of the data assimilation problem.

E.1 JASON's 'Carbon' Problem

To identify sources of carbon, or for that matter any atmospheric tracer gas, one envisions solving the linear transport equation for the concentration $c(x, y, z, t) = c(\mathbf{r}, t)$ of that constituent. This equation is formulated by Bocquet [1] for the ratio of the concentration and the density $\xi(\mathbf{r}, t) = \frac{c(\mathbf{r},t)}{\rho(\mathbf{r},t)}$

$$\frac{\partial \xi(\mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{r}} \cdot (\mathbf{u}(\mathbf{r}, t)\xi(\mathbf{r}, t)) = \mathbf{K}\nabla^2 \xi(\mathbf{r}, t) + s(\mathbf{r}, t), \tag{E-20}$$

but, as discussed by Bocquet, one may neglect local changes in the density as one follows a minor tracer constituent, and take $\rho(\mathbf{r}, t) =$ constant. Then we will be using

$$\frac{\partial c(\mathbf{r}, t)}{\partial t} + \mathbf{u}(\mathbf{r}, t) \cdot \nabla_{\mathbf{r}} c(\mathbf{r}, t) = \mathbf{K}\nabla^2 c(\mathbf{r}, t) + s(\mathbf{r}, t), \tag{E-21}$$

to characterize the transport of the tracer concentration from spatially distributed and time dependent sources $s(\mathbf{r}, t)$. $\mathbf{K}$ is a diffusion constant, and

$\mathbf{u}(\mathbf{r},t)$ is the local wind velocity; $\nabla_{\mathbf{r}} \cdot \mathbf{u}(\mathbf{r},t) = 0$ The idea is that with sufficient observations of the concentration, we should be able, assuming we know the wind field $\mathbf{u}(\mathbf{r},t)$, to identify a source term $s(\mathbf{r},t)$ that gives rise to the observed concentration field. Given the deterministic, or stochastic wind, $\mathbf{u}(\mathbf{r},t)$ this is a linear problem.

The full problem, as there are circumstances where the constituent can affect local radiative transfer, thus through local heating or cooling the local temperature, and thus through the equation of state the driving forces (e.g., pressure gradients) the local velocity field, namely, the wind, is more complex and, because of the advection in the Navier-Stokes equation

$$\frac{\partial \mathbf{u}(\mathbf{r},t)}{\partial t} + \mathbf{u}(\mathbf{r},t) \cdot \nabla \mathbf{u}(\mathbf{r},t) = -\frac{\nabla p(\mathbf{c}(\mathbf{r},t), T(\mathbf{r},t), ...)}{\rho_0} + \nu \nabla^2 \mathbf{u}(\mathbf{r},t) + \mathbf{F}^{ext}(\mathbf{r},t), \tag{E-22}$$

is nonlinear. Temperature transport, energy conservation,

$$\frac{\partial T(\mathbf{r},t)}{\partial t} + \mathbf{u}(\mathbf{r},t) \cdot \nabla T(\mathbf{r},t) = \kappa \nabla^2 T(\mathbf{r},t) + S(\mathbf{r},t), \tag{E-23}$$

adds additional nonlinearity. The full problem requires tracing the wind, through observations, the temperature, through observations, and other fields (entropy, salinity, ...) if necessary.

We will give a formal structure for using observations on $c(\mathbf{r},t), T(\mathbf{r},t)$, and $\mathbf{u}(\mathbf{r},t)$ to determine parameters in these equations: transport coefficients, sources, constants in the equations of state, ... for the full nonlinear problem, and we will discuss how to implement this formal structure when the wind is given, either deterministically or as a known Gaussian distribution about a mean wind, for the concentration of a tracer gas (we call it carbon) to estimate the fixed parameters, the sources, and the values of the concentration $c(\mathbf{r},T)$ at the end of an observation period $\{t_0, ..., T\}$.

E.2 Introduction

Assimilating the information in observed data into models of a dynamical system when there are errors in the measurements, errors in the models, and

uncertainty about the precise state of a system when the assimilation process begins has stimulated discussions about placing data assimilation (DA) in a probabilistic setting [2, 3, 4]. The setting we have in mind is the following: a physical system is observed and a D-dimensional model of the system is developed based on some relevant principles. How the model is developed is not part of the DA problem, but we want to use DA to establish various aspects of the dynamical rules including estimates of any unknown fixed parameters in the model formulation, estimates of the state variables which are not observed, and estimates of any external forcings of the dynamics.

Prediction using a model requires information, preferably accurate, on any fixed parameters in the model, and on the state variables of the model $\mathbf{x}(T)$ when prediction begins. In linear models this is a challenge, and in nonlinear models it has difficulties associated with intrinsic instabilities in the model that must be addressed. There is, interestingly, a silver lining to this. One issue, not generally addressed in the literature, is the question of how many observations one must make to enable the accurate estimation of the state variables $\mathbf{x}(T)$ and the parameters. There are two aspects to this question: (1) how often must observations be made, and (2) how many observations are required at each time a set of measurements are performed. We comment on both.

To define notation, we review the formulation of DA. Most systems of interest are described by model partial differential equations (PDEs) describing fields $\phi(x, y, z, t) = \phi(\mathbf{r}, t)$ satisfying equations of the form

$$\frac{\partial \phi(\mathbf{r}, t)}{\partial t} = \mathbf{F}(\phi(\mathbf{r}, t), \nabla_{\mathbf{r}} \phi(\mathbf{r}, t), \mathbf{p}, \mathbf{s}(\mathbf{r}, t)), \tag{E-24}$$

where $\mathbf{F}$ describes the linear or nonlinear dynamics of the dynamical variables $\phi(\mathbf{r}, t)$, $\mathbf{p}$ are a set of fixed parameters, typically transport coefficients, conductivities, densities, etc, and $\mathbf{s}(\mathbf{r}, t)$ are external forces driving the dynamics of the fields $\phi(\mathbf{r}, t)$.

These equations are now discretized in both space and time to represent the observation points and to provide a framework for numerical solution of the

dynamical problem. Any indices on the fields $\phi(\mathbf{r}, t)$ indicating components of vectors (velocity, for example, or components of the overall degrees of freedom), and the spatial labels in the discretization of $\mathbf{r} = (x, y, z)$ are incorprated into a label $a = 1, 2, ..., D$. The model degrees of freedom are now called $x_a(t)$ and the PDEs have become a D-dimensional set of ODEs:

$$\frac{dx_a(t)}{dt} = F_a(\mathbf{x}(t), \mathbf{p}, \mathbf{s}(t)). \tag{E-25}$$

These equations are now solved numerically in discrete time. We label time as t_n or indicate time by the integer index n, and in this way we change our differential equations to discrete time maps relating the state of the system at time t_n, $\mathbf{x}(n)$, to the state of the system one step later $\mathbf{x}(n+1)$. The numerical scheme for solving the differential equations can be this discrete time representation. Generally it is an implicit relation of the form

$$g_a(\mathbf{x}(n+1), \mathbf{x}(n), \mathbf{p}, \mathbf{s}(n)) = 0;\ a = 1, 2, ..., D. \tag{E-26}$$

If we can arrange for the discrete time dynamics to be explicit, we can write this in the equivalent form

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{p}, \mathbf{s}(n)), \tag{E-27}$$

and this dynamical rule may be linear in the $\mathbf{x}(n)$ or nonlinear in the $\mathbf{x}(n)$. In this formulation we see that the forcings are exactly on the same footing as the fixed parameters–i.e. there is no differential equation for them–except that there are many quantities $\mathbf{s}(n)$ which we must estimate based on the dynamical rule and the measurements. The boundary conditions for the PDEs enter the discrete time formulation as constraints on the solutions involving further parameters, all of which are added to the $\mathbf{p}$.

Measurements are now made at discrete times within a temporal window $\{t_0, t_1, ..., t_m = T\}$. We call these measurements $y_l(n);\ l = 1, 2, ..., L$, and typically $L < D$, often $L << D$.

Further, we must recognize that the measurements are noisy, the model has errors, and the state of the system at the initiation of measurements t_0, namely $\mathbf{x}(0)$, is known only with uncertainty.

To connect the model dynamics and the measurements, we must make those measurements a further part of our model and relate the L observations at each time $\{t_0, t_1, ..., t_m = T\}$ to the state of the model. In this way we associate the measurement functions $h_l(\mathbf{x}(n));\ l = 1, 2, ..., L$ with the observed quantities as known functions of the model dynamical variables. In estimating quantities associated with the model, states and parameters, we want to adjust the model so its output $\mathbf{x}(n)$ is such that at any observation time $h_l(\mathbf{x}(n))$ matches, using some comparison metric, the observations $y_l(n)$. The fact that $L < D$ means we are faced with what control theorists for many decades have called the 'observer' problem [5]: we have $L < D$ measurements on a system described by a D-dimensional dynamical system; how large must L be to allow estimation of all D components of the model state $x_a(t)$ and all the parameters in the model. If $L = D$, as might be the case, we need only estimate the parameters, for then the observations give us $\mathbf{x}(T)$ at the end of the observation period, and we can then use the dynamical rule $\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{p}, \mathbf{s}(n))$ to predict forward in time.

The overall model is now set: it consists of the dynamics for the model state variables $\mathbf{x}(n)$, the connection of the measurements $y_l(n)$ with the model states $\mathbf{x}(n)$: $y_l(n) \approx h_l(\mathbf{x}(n))$, and the set of parameters $\mathbf{p}$ comprising fixed parameters in the dynamical rule $\mathbf{x}(n) \rightarrow \mathbf{x}(n+1)$, in the 'measurement functions' $h_l(\mathbf{x})$, in the boundary conditions in the original PDEs, and in the parametrized forcing $\mathbf{s}(n)$.

As a shorthand notation we will now introduce D coordinates which emphasize the measurements $\mathbf{z} = [z_1, z_2, .., z_D] = [h_1(\mathbf{x}), h_2(\mathbf{x}), ..., h_L(\mathbf{x}), x_{L+1}, ...x_D]$. Our task is to use the observations to produce an estimate of the conditional probability for the state being in the location $\mathbf{z}(t_m) = \mathbf{z}(m)$ at the end of an observation window $\{t_0, t_1, ..., t_m = T\}$, conditioned on the observations $\mathbf{Y}(m) = \{\mathbf{y}(m), \mathbf{y}(m-1), ..., \mathbf{y}(0)\}$. We call this $P(\mathbf{z}(m)|\mathbf{Y}(m))$.

The First Step

The discussion of this conditional probability goes back to Jazwinski [6, 4] or

further. There are two steps: first we use Bayes' rule to relate the conditional probability $P(\mathbf{z}(m)|\mathbf{Y}(m))$ utilizing $m+1$ measurements to the conditional probability $P(\mathbf{z}(m)|\mathbf{Y}(m-1))$ utilizing m measurements:

$$P(\mathbf{z}(m)|\mathbf{Y}(m)) = \left\{ \frac{P(\mathbf{z}(m), \mathbf{y}(m)|\mathbf{Y}(m-1))}{P(\mathbf{z}(m)|\mathbf{Y}(m-1))\, P(\mathbf{y}(m)|\mathbf{Y}(m-1))} \right\} P(\mathbf{z}(m)|\mathbf{Y}(m-1)), \tag{E-28}$$

This is an identity on conditional probabilities. It is usually written as [4]

$$P(\mathbf{z}(m)|\mathbf{Y}(m)) = \left[\frac{P(\mathbf{y}(m)|\mathbf{z}(m), \mathbf{Y}(m-1))}{P(\mathbf{y}(m)|\mathbf{Y}(m-1))} \right] P(\mathbf{z}(m)|\mathbf{Y}(m-1)), \tag{E-29}$$

which while correct, does not emphasize the factor

$$\left\{ \frac{P(\mathbf{z}(m), \mathbf{y}(m)|\mathbf{Y}(m-1))}{P(\mathbf{z}(m)|\mathbf{Y}(m-1))\, P(\mathbf{y}(m)|\mathbf{Y}(m-1))} \right\}, \tag{E-30}$$

as the exponential of the **conditional mutual information**, in the sense of Shannon, between the L-dimensional measurement $\mathbf{y}(m)$ at time t_m and the D-dimensional state $\mathbf{z}(m)$ at the same time, conditioned on the earlier measurements $\mathbf{Y}(m-1)$:

$$\begin{aligned} & \exp\left[\log\left\{ \frac{P(\mathbf{y}(m)|\mathbf{z}(m), \mathbf{Y}(m-1))}{P(\mathbf{y}(m)|\mathbf{Y}(m-1))} \right\} \right] \\ = \;& \exp\Big[MI(\mathbf{z}(m), \mathbf{y}(m)|\mathbf{Y}(m-1)) \Big]. \end{aligned} \tag{E-31}$$

In a certain practical sense, this difference is not important, but Equation (E-31) does give a precise information theoretic meaning to the addition of one more measurement, the one at t_m, namely $y(m)$, to the knowledge we have of the conditional probability distribution of the state of the model at t_m. To see this write from Equation (E-28)

$$\begin{aligned} \log\left\{ \frac{P(\mathbf{z}(m)|\mathbf{Y}(m))}{P(\mathbf{z}(m))} \right\} \Big\} \;\; &= \\ MI(\mathbf{z}(m), \mathbf{y}(m)|\mathbf{Y}(m-1)) \;\; &+ \;\; \log\left\{ \frac{P(\mathbf{z}(m)|\mathbf{Y}(m-1))}{P(\mathbf{z}(m))} \right\}, \end{aligned} \tag{E-32}$$

showing that the added information (in nats) arising from one additional measurement is precisely the conditional mutual information. This suggests

the view of a communication channel unidirectionally connecting the transmitter (the data source) and the receiver (the model) into which one can put additional information by making a new measurement. This last result tells us how much information about the measurement is transmitted through the channel connecting the data source and the model.

The Second Step

All that said, we now can take the required second step. To relate the conditional distribution $P(\mathbf{z}(m)|\mathbf{Y}(m-1))$ to $P(\mathbf{z}(m-1)|\mathbf{Y}(m-1))$ and then have a recursion relation between $P(\mathbf{z}(m)|\mathbf{Y}(m))$ and $P(\mathbf{z}(m-1)|\mathbf{Y}(m-1))$ to allow the determination of $P(\mathbf{z}(m)|\mathbf{Y}(m))$ from $P(\mathbf{z}(0))$, we use the Chapman-Kolmogorov relation

$$P(\mathbf{z}(m)|\mathbf{Y}(m-1)) = \int d^D z(m-1) P(\mathbf{z}(m)|\mathbf{z}(m-1)) P(\mathbf{z}(m-1)|\mathbf{Y}(m-1)), \tag{E-33}$$

which is true for dynamical rules giving $\mathbf{z}(n+1)$ from knowledge of $\mathbf{z}(n)$. We may also express this in terms of the original state variables $\mathbf{x}(n)$ by noting

$$\begin{aligned} d^D z(n) &= \frac{\partial(\mathbf{z}(n))}{\partial(\mathbf{x}(n))} d^D x(n), \\ \frac{\partial(\mathbf{z})}{\partial(\mathbf{x})} &= J_L(\mathbf{x}) = \frac{\partial(h_1(\mathbf{x}), h_2(\mathbf{x}), ..., h_L(\mathbf{x}))}{\partial(x_1, x_2, ..., x_L)} \end{aligned} \tag{E-34}$$

which is an $L \times L$ matrix of derivatives of the measurement functions with respect to $\{x_1, x_2, ..., x_L\}$.

The transition probability $P(\mathbf{z}(n+1)|\mathbf{z}(n))$ is determined by the dynamical rule. If the dynamics is deterministic, then $\mathbf{z}(n+1) = \mathbf{q}(\mathbf{z}(n), \mathbf{p})$, and

$$P(\mathbf{z}(n+1)|\mathbf{z}(n)) = \delta^D(\mathbf{z}(n+1) - \mathbf{q}(\mathbf{z}(n), \mathbf{p})), \tag{E-35}$$

while if there is model error or finite resolution, the sharp distribution of a delta function is broadened by some measure of that resolution or error. Often this might be represented by additive noise, but it may be more general.

Now take the recursion relation:

$$P(\mathbf{z}(m)|\mathbf{Y}(m)) = \exp\Big[MI(\mathbf{z}(m),\mathbf{y}(m)|\mathbf{Y}(m-1))\Big]$$
$$\int d^D z(m-1) P(\mathbf{z}(m)|\mathbf{z}(m-1)) P(\mathbf{z}(m-1)|\mathbf{Y}(m-1)), \quad \text{(E-36)}$$

and iterate it back from t_m to t_0 to find, with $\mathbf{X} = \{\mathbf{x}(m), \mathbf{x}(m-1), ..., \mathbf{x}(0)\}$,

$$\begin{aligned} P(\mathbf{z}(m)|\mathbf{Y}(m)) &= \int \prod_{n=0}^{m-1} d^D x(n) J_L(\mathbf{x}(n) e^{[-A_0(\mathbf{X}|\mathbf{Y})]}, \\ &= \int d\mathbf{X} e^{[-A_0(\mathbf{X}|\mathbf{Y})]}, \\ &= \int \prod_{n=0}^{m-1} d^D z(n) e^{[-A_0(\mathbf{X}|\mathbf{Y})]} \\ &= \int d\mathbf{Z} e^{[-A_0(\mathbf{X}|\mathbf{Y})]}, \end{aligned} \quad \text{(E-37)}$$

with

$$\begin{aligned} -A_0(\mathbf{X}|\mathbf{Y}) &= \sum_{n=0}^{m} MI(\mathbf{z}(n),\mathbf{y}(n)|\mathbf{Y}(n-1)) \\ &+ \sum_{n=0}^{m-1} \log[P(\mathbf{z}(n+1)|\mathbf{z}(n)] + \log[P(\mathbf{z}(0))]. \end{aligned} \quad \text{(E-38)}$$

This is a path integral over paths in $\mathbf{z}$ (or $\mathbf{x}$) space through a time window $\{t_0, t_1, ..., t_{m-1}\}$, namely a mD-dimensional integral.

We can now go back to $\mathbf{x}$ space by noting that probability densities in $\mathbf{z}$ are related to their counterparts in $\mathbf{x}$ via

$$P(\mathbf{z}) d^D z = P(\mathbf{x}) d^D x, \quad \text{(E-39)}$$

or

$$P(\mathbf{z}) = P(\mathbf{x}) \frac{(\partial \mathbf{x})}{(\partial \mathbf{z})}, \quad \text{(E-40)}$$

which involves the inverse of the Jacobian encountered above. Sorting out the various Jacobians leads to an expression for the quantities of real interest in data assimilation, the conditional expected values of functions of state

variables. If a general variable on the path $\mathbf{X} = \{\mathbf{x}(m), \mathbf{x}(m-1), ..., \mathbf{x}(0)\}$ through the observation window is written as $G(\mathbf{X})$, we have now

$$< G(\mathbf{X}) >= E[G(\mathbf{X})|\mathbf{Y}] = \frac{\int d\mathbf{X} G(\mathbf{X}) e^{[-A_0(\mathbf{X}|\mathbf{Y})]}}{\int d\mathbf{X} e^{[-A_0(\mathbf{X}|\mathbf{Y})]}}, \qquad \text{(E-41)}$$

with

$$\begin{aligned} -A_0(\mathbf{X}|\mathbf{Y}) &= \sum_{n=0}^{m} MI(\mathbf{z}(\mathbf{x}(n)), \mathbf{y}(n)|\mathbf{Y}(n-1)) \\ &+ \sum_{n=0}^{m-1} \log[P(\mathbf{x}(n+1)|\mathbf{x}(n)] + \log[P(\mathbf{x}(0))]. \end{aligned} \qquad \text{(E-42)}$$

This tells us that the only place the observation function enters the evaluation of physical quantities of interest is in the mutual in formation term of the action. There it appears as $\mathbf{h}(\mathbf{x})$. No inversions or 'retrievals' are required to evaluate $< G(\mathbf{X}) >$ for any quantity of interest.

The path integral formulation is found in [7] with examples of the direct Monte Carlo integration of the high dimensional integral for a model of some geophysical interest found in [8] and the trickiness of a saddle path evaluation of the integral discussed in [9].

E.3 Saddle Path Approximation: 4DVar

A natural and familiar way to approximate integrals such as $(E-41)$ is to expand the integrand, here $A_0(X|Y)$ about a minimum path $\mathbf{S}$

$$\begin{aligned} &A_0(\mathbf{X}|\mathbf{Y}) = A_0(\mathbf{S}|\mathbf{Y}) + \frac{1}{2}(\mathbf{X}-\mathbf{S})\frac{\partial A_0(\mathbf{X}|\mathbf{Y})}{\partial \mathbf{X}\,\partial \mathbf{X}}|_{\mathbf{X}=\mathbf{S}}(\mathbf{X}-\mathbf{S}) + ... \\ &\frac{\partial A_0(\mathbf{X}|\mathbf{Y})}{\partial \mathbf{X}}|_{\mathbf{X}=\mathbf{S}} = 0. \end{aligned} \qquad \text{(E-43)}$$

and evaluate the resulting Gaussian integral to establish corrections to the path $\mathbf{S}$. This is known as (weak) 4DVar in the geophysics literature. The 'weak' means the equations of motion are satisfied in an average way rather than precisely as in a deterministic ('strong') manner.

Unfortunately, because of intrinsic instabilities in chaotic flows of nonlinear systems, the procedure of searching for the path $\mathbf{S}$ encounters impediments in the form of numerous local minima making the search unfeasible. Regularizing the search can cure this [9].

In the linear problem we address for the tracer problem with specific wind field, this should not be an issue.

E.4 How Many Measurements are Required?

Implicit in many discussions of the tracer source reconstruction problem, one assumes more measurements would always be better. From discussions of observer theory [5] one knows that this might not actually be necessary. By this one means that in some problems it may be that parameters and unobserved state variables may be estimated when L reaches a certain value ($L < D$ still) and no more may be required. A linear example may serve to illustrate this. Suppose the original, data generating dynamics is $\mathbf{w}(n+1) = \mathbf{M} \cdot \mathbf{w}(n)$. $\mathbf{M}$ may have eigenvalues outside the unit circle. The 'model' system is taken to be $\mathbf{x}(n+1) = \mathbf{M} \cdot \mathbf{x}(n)$ and we want to construct an 'observer' system allowing some measurements of the $\mathbf{w}$ allowing an estimate of all the the $\mathbf{x}(n)$ so $|\mathbf{x}(n) - \mathbf{w}(n)| \to 0$ as time goes by.

One such construction, called a Luenberger observer, is to add to the model dynamics a term

$$\mathbf{x}(n+1) - \mathbf{M} \cdot \mathbf{x}(n) + \mathbf{K}(\mathbf{w}(n) - \mathbf{x}(n)), \tag{E-44}$$

where $\mathbf{K}$ is a constant matrix of dimension $D \times L$ and observations of L states $\mathbf{w}$ are available at each time n. The equation for $\mathbf{w}(n) - \mathbf{x}(n) = \Delta(n)$ is

$$\Delta(n+1) = (\mathbf{M} - \mathbf{K})(\Delta(n)), \tag{E-45}$$

and for large n we see that $\Delta \to 0$ if the eigenvalues of the matrix $\mathbf{M} - \mathbf{K}$ lie within the unit circle, regardless of where the eigenvalues of $\mathbf{M}$ may sit.

So, even if the dynamics $\mathbf{w}(n) \rightarrow \mathbf{M} \cdot \mathbf{w}(n)$ is *unstable* one may construct an observer to estimate the unobserved states.

In the case the dynamics is nonlinear, and the analog of the matrix $\mathbf{M}$ is time dependent, a similar construction can be used to cure the instabilities on the submanifold of the data generation plus model space where $\mathbf{w}(n) \approx \mathbf{x}(n)$, the synchronization' manifold.

In the language of the path integral this takes two forms. In the case of doing a saddle path integral the ability to find a path $\mathbf{S}$

$$\frac{\partial A_0(\mathbf{X}|\mathbf{Y})}{\partial \mathbf{X}}|_{\mathbf{X}=\mathbf{S}} = 0, \tag{E-46}$$

subject to the dynamics, deterministic probably, reflects the instability on the synchronization manifold. It does this by expressing a large number of local minima connected with the incoherence of the data signal $y_l(t_n)$ and the model representation of this $h_l(\mathbf{x}(t_n))$ when both are chaotic but not synchronized. The role of the analog of the matrix $\mathbf{K}$, the observer, is to synchronize the data and the model output. This eliminates the local instabilities.

In this context, then, the minimum number of required measurements to allow the estimation of the parameters and unobserved states is the number of independent instabilities on the synchronization manifold. This is cast into a statement about the indices of such instabilities, the conditional Lyapunov exponents (CLEs). So, given a model, and given data either generated by the model or by a data source, evaluae the CLEs of the model given the data, and determine how many independent data time series are required to force all CLEs to be negative. That determines L.

There is another way, which is not as precise, but probably equivalent. In doing the path integral, say by Monte Carlo methods, one will be successful if there are a few, even better only one, paths that dominate the contributions to the integral. In the classical case here, this has to do with minima of the action $A_0(\mathbf{X}|\mathbf{Y})$–so not unrelated to the saddle path approximation. If

one has data that visits the minima of the action, the main contributions to the integral will be picked out, and with luck other paths will contribute an exponentially small quantity to the integral.

In any case, the quantity $\exp[-A_0(\mathbf{X}|\mathbf{Y})]$ is the density of paths in the path space of the $\mathbf{X}$–$(m+1)D$ dimensional space. The same density of paths is generated by solutions to the Langévin equation expressed in (fake) 'time' s:

$$\frac{d\mathbf{X}(s)}{ds} = -\frac{1}{2}\frac{\partial A_0(\mathbf{X}(s)|\mathbf{Y})}{\partial \mathbf{X}(s)} + \eta(s), \tag{E-47}$$

where $\eta(s)$ is Gaussian white noise of zero mean and variance unity.

So one way to explore the structure of the dependence of $A_0(\mathbf{X}|\mathbf{Y})$ in $\mathbf{X}$ space is to solve the Langévin equation for a collection of initial conditions $\mathbf{Y}^j(s=0;\ j=1,2,...,$ and see how many different values of $A_0(\mathbf{X}(s\to\infty|\mathbf{Y})$ are observed. In practice, one simply takes s large, not ∞, of course, and looks at the values of the action that arise. In simple examples one does this for zero measurements, one measurement, two measurements, ... For zero measurements, one typically finds, for nonlinear dynamics of the model, many minima of the action, corresponding to the many local minima in the saddle path method. As one adds measurements, the number of minima decreases and then settles into one or a few at just about the same value of L as the CLE estimation above would indicate.

The upshot of all this is that given a model, reflecting the resolution one wishes to accomplish in space and time, one can use this feature to estimate how many measurements at any given time one should make. If one has this minimum number of measurements at each observation time, then by evaluating the expected path **and** the RMS variation about that path over the whole observation window, one can now estimate the accuracy with which the desired measurement $\mathbf{x}(t_m = T)$ can be evaluated. This accuracy should certainly decrease as one takes away measurements at each time within the observation window, and this may indicate, given a desired level of accuracy, how many measurements fewer than the 'required' number will be acceptable.

Finally, and this is another question, one may ask how often measurements need to be made within the observation window. So assuming one makes the required L measurements at each time a set of observations are made, what happend if the time between measurements is increased? The rule of thumb here comes from noting that between measurements one moves the model system forward using the model dynamics. The predictability of those dynamics is associated with the Lyapnov exponents of the model system wherein errors in the knowledge of a state at time t_n $\Delta\mathbf{x}(n)$ grows exponentially rapidly: $\Delta\mathbf{x}(t_k \geq t_n) \approx \Delta\mathbf{x}(t_n)\exp[\lambda(t_k - t_n)]$. If measurements are made with time intervals between measurements longer than $1/\lambda$, then the errors in knowledge of the system state are likely to be too big to recover with another measurement. Stated otherwise, the coherence of the state after a few times $1/\lambda$ with the earlier state is so low that measurements made so infrequently are unlikely to be able to efficiently or accurately redirect the state to the correct region of phase space from where accurate predictions may be accomplished.

E.5 Other Approaches to Data Assimilation

Assimilating information from data into models and estimates of states and parameters has many approaches which can be formulated as ways to approximate the path integral.

- 4DVar

  The saddle path approximation to the path integral requires a search for a path or paths $\mathbf{S}$ satisfying

  $$\frac{\partial A_0(\mathbf{X}|\mathbf{Y})}{\partial \mathbf{X}}|_{\mathbf{X}=\mathbf{S}} = 0, \tag{E-48}$$

  and this is often seen as a maximum likelihood estimation. This method is used in the incorporation of observations by the European Center for Medium Range Weather Forecasts as well as in modeling of Co_2 transport [10]. The outcome from this kind of calculation is a path $\mathbf{S}$

through the assimilation or observation window, and the state at the end of the window $\mathbf{s}(T) \approx \mathbf{x}(T)$ and the estimates of fixed parameters using the information transferred in the window are to be used as inputs to the underlying model for predicting for $t > T$. This approach does not give a method for estimating the RMS error or other moments about the 'optimal' path $\mathbf{S}$.

The path integral, as noted, does that by drawing attention to the Gaussian integral about the path $\mathbf{S}$.

- Shadowing Paths

  This is a method [11] that focuses on the satisfaction of the model dynamical equations $\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{p}))$. It minimizes $\frac{1}{2}\sum_{n=0}^{m-1}(\mathbf{x}(n+1) - \mathbf{f}(\mathbf{x}(n), \mathbf{p}) - \xi(n))^2$ starting with observations for $\mathbf{X}$ across the measurement window $\{t_0, t_1, ..., t_m\}$ and $\xi = 0$. Moves to minimize this are done with prescribed shifts in $\mathbf{X}$ and ξ. This is equivalent to minimizing the path integral emphasizing the dynamics term starting with full satisfaction of the measurement error term, resulting in a single 'optimal' path. As in the evaluation of the path integral, all the information within the measurement window is used. However, all state variables must be measured, and there is no estimate of the corrections to the 'optimal' path that results. It is not an 'ensemble' method. Again those restrictions are lifted when the full path integral is used.

- Ensemble Filters

  This is a method very close to the ideas in the path integral approach. Here one tries to satisfy the two step process described earlier by using the dynamical equations to carry out the $\mathbf{x}(n) \rightarrow \mathbf{x}(n+1)$ step, then using an assumption about the distribution of state variables after the operation of the dynamics, information from the measurements is incorporated into knowledge of the probability distribution at the next time step.

  In particular, for what is called the Ensemble Kalman Filter [12] or its local (in space) version [13], one assumes that at each time in the measurement window the distribution of states $P(\mathbf{x}(n-1))$, conditioned

on earlier measurements is Gaussian. At t_n, before the dynamics acts, one uses what knowledge there is of $P(\mathbf{x}(n-1)$ and from that evaluates a mean and a covariance matrix. Points are moved one step in time by the nonlinear dynamics of the model, and the new distribution, not Gaussian at all, is approximated by a new Gaussian $P(\mathbf{x}(n))$. The mean and covariance of the new Gaussian is adjusted using the rules of a Kalman filter, appropriate to a linear optimization problem, to the next state, where the adjusted, still Gaussian, distribution is operated on by the nonlinear dynamics, giving a distribution which is approximated by a Gaussian, etc. In the *local* ensemble Kalman filter strategy, the distribution after the action of the nonlinear model dynamics, one uses a collection of local Gaussians in the spirit of Anderson and the spirit of local radial basis functions used as a collective sum of Gaussians to approximate the actual, non-Gaussian, distribution. When the information from the measurements is transmitted to the model and the state variable distribution is approximated by a set of local (in space) Gaussians, one is no longer doing a linear approximation anywhere but a sequence of piece-wise linear (in space) approximations, and this can do a very credible job of approximating quite complex distributions $P(\mathbf{x}(n))$ even in high dimensions. This, presumably, underlies the success of the *local* ensemble filter.

Because this is an ensemble filter, one has information both about an 'optimal' path through the assimilation window as well as nearby paths, so sample covariance matrices and higher moments about the mean can be evaluated.

This local ensemble Kalman filter is what has been used by Fung, Kalnay and collaborators in combining the transport equation for CO_2 concentration with meterological models to determine the wind field appearing in the transport equation.

E.6 Relation to the Carbon Problem

Deterministic Wind

Following Bocquet [1] we are interested in reconstructing the source term $s(\mathbf{r},t)$ for the production and transport of a tracer whose concentration $c(\mathbf{r},t)$ is governed by the linear equation

$$\frac{\partial c(\mathbf{r},t)}{\partial t} + \mathbf{u}(\mathbf{r},t)\cdot\nabla_{\mathbf{r}}c(\mathbf{r},t) - \nabla_{\mathbf{r}}\cdot(\mathbf{K}\nabla_{\mathbf{r}}c(\mathbf{r},t)) = s(\mathbf{r},t), \tag{E-49}$$

and L observations are made

$$\mathbf{y}_l = \int d^3r dt \pi_l(\mathbf{r},t)c(\mathbf{r},t), \tag{E-50}$$

where the $\pi_l(\mathbf{r},t)$ are a normalized weighting factors for the importance of measurements at time t, in the interval $[t_0,T]$ over some region in $\mathbf{r}=(x,y,z)$ space.

The question is how to estimate $\mathbf{s}(\mathbf{r},t)$ from these dynamics and these measurements. First we will assume that the relation between the L measurements y_l and the concentrations $c(\mathbf{r},t)$ comes from the projection function of Bocquet [1] measurements sampling the concentration at L spatial locations at $m+1$ times $\{t_0,t_1,...,t_m\}$ then

$$H_{la}c(\mathbf{r}_a,t_n) = H_{la}c_a(n), \tag{E-51}$$

where the $L\times D$ measurement matrix is a summary of Bocquet's projection operator $p(\mathbf{r},t)$. The linear PDE for the concentration $c(\mathbf{r},t)$ is discretized in space at D grid points and turned into a map from time t_n to time t_{n+1} of the form

$$\begin{aligned}\frac{c_a(n+1)-c_a(n)}{\tau} &+ \sum_{k=1}^{3} u_k(\mathbf{r}_a,n)[c(\mathbf{r}_{a+1},n)-c(\mathbf{r}_a,n)] \\ &- K[c(\mathbf{r}_{a+1},n)-2c(\mathbf{r}_a,n)+c(\mathbf{r}_{a-1},n)] - s_a(n) = 0;\end{aligned}$$

or

$$c_a(n+1) = \sum_{b=1}^{D} M_{ab}(n)c_b(n) + s_a(n). \tag{E-52}$$

To do the path integral we must make some assumptions on the way noise enters the measurements and errors spoil the infinite resolution of the model dynamics. We assume that the noise is Gaussian at each measurement time and independent at each measurement time. The conditional mutual information term then becomes

$$\frac{R_m}{2}\sum_{n=0}^{m}\sum_{l=1}^{L}(y_l(n)-\sum_{a=1}^{D}H_{la}c_a(n))^2, \tag{E-53}$$

and $\frac{1}{R_m}$ is proportional to the variance of the measurement noise, taken to be the same for each measurement.

If we broaden the deterministic, discrete in time and space, version of the dynamics by an assumption about the reduced resolution of the model **or** fluctuations in the source term, we can take

$$P(\mathbf{c}(n+1)|\mathbf{c}(n)) = \delta^D(\mathbf{c}(n+1)-\mathbf{M}\mathbf{c}(n)-\mathbf{s}(n)) \tag{E-54}$$

into

$$P(\mathbf{c}(n+1)|\mathbf{c}(n)) = \exp[-\sum_{n=0}^{m-1}(\mathbf{c}(n+1)-\mathbf{M}\mathbf{c}(n)-\mathbf{s}(n))^2/2\sigma_f^2], \tag{E-55}$$

where σ_f is the standard deviation of the error or fluctuations in the model. We might, as below, attribute these to fluctuations in the source; we might not.

With these assumptions, we may write the 'action' $A_0(\mathbf{C}|\mathbf{Y})$, where $\mathbf{C} = \{\mathbf{c}(m),\mathbf{c}(m-1),...,\mathbf{c}(0)\}$, as

$$\begin{aligned} A_0(\mathbf{C}|\mathbf{Y}) &= \frac{R_m}{2}\sum_{n=0}^{m}\sum_{l=1}^{L}(y_l(n)-\sum_{a=1}^{D}H_{la}c_a(n))^2 \\ &+ \frac{R_f}{2}\sum_{n=0}^{m-1}(\mathbf{c}(n+1)-\mathbf{M}\mathbf{c}(n)-\mathbf{s}(n))^2, \end{aligned} \tag{E-56}$$

Using this action we can calculate the moment generating function $\chi(K)$; $\mathbf{K} = \{k_a(m),k_a(m-1),...,k_a(0)\}$

$$\exp[\chi(\mathbf{K})] = \int d\mathbf{C}\exp[-A_0(\mathbf{C}|\mathbf{Y})+\mathbf{K}\cdot\mathbf{C}], \tag{E-57}$$

by noting that $A_0(\mathbf{C}|\mathbf{Y}) = \frac{1}{2}\mathbf{C}^T\mathbf{Q}(\mathbf{M}, R_m, R_f, \mathbf{Y}, \mathbf{H})\mathbf{C} - \mathbf{W}(\mathbf{M}, R_m, R_f, \mathbf{s}) \cdot \mathbf{C}$ + terms independent of $\mathbf{C}$, we can perform the Gaussian integral over $\mathbf{C}$ to find

$$\chi(K) = \frac{1}{2}\mathbf{W} \cdot \mathbf{Q}^{-1}\mathbf{W} + \mathbf{W} \cdot \mathbf{Q}^{-1}\mathbf{K} + \frac{1}{2}\mathbf{K} \cdot \mathbf{Q}^{-1}\mathbf{K} - \frac{1}{2}\text{trace}\log[\frac{\mathbf{Q}}{2\pi}]. \quad \text{(E-58)}$$

This ignores some constants independent of $\mathbf{K}$. The dependence on the source $s_a(n)$ is in the vector $\mathbf{W}$ only. There is a quadratic term in $s_a(n)$ but it cancels out in determining the expectation value of $c_a(n)$. The matrix $\mathbf{M}$ is linear in the wind field $\mathbf{u}_a(n)$, and $\mathbf{u}(n)$ appears both in the vector $\mathbf{W}$ and in the matrix $\mathbf{Q}$.

The conditional expectation value, conditioned on the measurements, of the concentration at space location $\mathbf{r}_a$ and time t_n, namely $< c_a(n) >$ is

$$< c_a(n) >= \frac{\partial\chi(\mathbf{K})}{\partial k_a(n)}|_{\mathbf{K}=0} = \mathbf{Q}^{-1} \cdot \mathbf{W}, \quad \text{(E-59)}$$

and the covariance about the mean is

$$< c_a(n)c_b(n') > - < c_a(n) >< c_b(n') >= \frac{\partial^2\chi(K)}{\partial k_a(n)\,\partial k_b(n')}|_{\mathbf{K}=0} = \mathbf{Q}^{-1}. \quad \text{(E-60)}$$

The expected value of the concentration is linear in the source, as expected from a linear problem, while the mean source value does not enter the covariance about the mean. It is in the vector $\mathbf{W}$ only.

However, if we attribute all fluctuations in the transport dynamics–we called them 'model errors' before–to fluctuations in the source term, so $\Delta s \propto \sigma_f$, then if we were to track the mean concentration in space (index a) and time, we could estimate the source term, and by tracking the covariance about that mean, we could estimate σ_f as well.

In this formulation of the problem, where we assume Gaussian noise in the measurements **and** Gaussian errors in the model totally attributed to source fluctuations, we can directly estimate both source strength and fluctuations from the concentration measurements. The relation between the actual measurements $y_l(n)$ and the concentrations requires inversion of the rectangular

$L \times D$ matrix H_{la}, and in this Gaussian context the inverse is naturally the pseudo-inverse $[\mathbf{H}^T \cdot \mathbf{H}]^{-1}\mathbf{H}$.

If we relax the assumptions about the distribution of noise n the measurements or the source of nondeterminism in the dynamics $\mathbf{c}(n+1) = \mathbf{M}\mathbf{c}(n) + \mathbf{s}(n)$, then we cannot do the integrals exactly, and we must perform the integrals, even in the linear dynamics case using some tool such as a Monte Carlo calculation.

Stochastic Wind

In a full nonlinear problem we would estimate the parameters such as K, the source locations as a function of time $s_a(n)$ and the wind $\mathbf{u}_a(n)$, and this involves the full nonlinear problem. Here, if the wind is given to us as a mean wind $\mathbf{u}_a^{(0)}(n)$ plus an error estimate in the form of a covariance matrix $R_{an,bn'} =< (\mathbf{u}_a(n) - \mathbf{u}_a^{(0)}(n))(\mathbf{u}_b(n') - \mathbf{u}_b^{(0)}(n') >$, we can find the effect of error in the wind on the estimation of the expected concentration $< c_a(n) >$ and the covariance about the mean $< c_a(n)c_b(n') > - < c_a(n) >< c_b(n') >$ by computing the averages

$$\int \prod_{n=0}^{m} d^3Du(n)\mathbf{Q}^{-1}\cdot\mathbf{W} \exp 1/2[(\mathbf{u}_a(n)-\mathbf{u}_a^{(0)}(n))R^{-1}_{an,bn'}(\mathbf{u}_b(n')-\mathbf{u}_b^{(0)}(n'))]\frac{1}{\det[2\pi\mathbf{R}]}$$

and

$$\int \prod_{n=0}^{m} d^3Du(n)\mathbf{Q}^{-1} \exp 1/2[(\mathbf{u}_a(n)-\mathbf{u}_a^{(0)}(n))R^{-1}_{an,bn'}(\mathbf{u}_b(n')-\mathbf{u}_b^{(0)}(n'))]\frac{1}{\det[2\pi\mathbf{R}]},$$

we will have an expression for the expected source from the first and the covariance of the concentration estimations from the second, averaged over the distribution of the wind.

A Tiny Example

To see how this works out in a tiny example, we select $D = 1$, namely concetration at one spatial point $c(n)$, $L = 1$, one measurement, $m = 1$, so two measurement times. The 'path' is comprised of the $(m+1)D = 2$ vector $\{c(0), c(1)\}$. We select the measurement function to have $H = 1$. The action

for the apth integral is

$$A_0(C|Y) = \frac{R_m}{2}((y(0)-c(0))^2+(y(1)-c(1))^2)+\frac{R_f}{2}[c(1)-(M_1+M_2u)c(0)-s(0)]^2, \tag{E-61}$$

where we made the 'wind' a scalar and recalled the linear dynamics from the transport equation is linear in the wind. Note the coefficient of u, is related to the gradient of the concentration.

The generating function for the mean $< c(0) >$ and $< c(1) >$ and the other moments is given by

$$\chi(k_0, k_1) = \log\left[\int dc(0)dc(1)\exp[-A_0(C|Y)]P(c(0))]\ \exp[k_0c(0) + k_1c(1)]\right], \tag{E-62}$$

where $P(c(0))$ is the distribution of concentration at the initial time. We will take this to be unity here, though if it is a Gaussian $\log[P(c(0))] \propto \frac{-(c(0)-c_0)^2}{2\sigma_c^2}$, that poses no special problem here.

The strategy is to integrate over $\mathbf{C} = \{c(0), c(1)\}$, holding the 'wind' u fixed, then integrate over the distribution of the wind. In itself that might come from the underlying turbulent fluid dynamics, but we do not address that too.

In our earlier notation, the integral to be done is

$$\int d^2C\ \exp[-1/2\mathbf{C}\cdot\mathbf{Q}\cdot\mathbf{C} + \mathbf{W}\cdot\mathbf{Q} + \mathbf{K}\cdot\mathbf{C}], \tag{E-63}$$

where $\mathbf{K} = \{k_0, k_1\}$, W is the two component vector $\mathbf{W} = \{-R_m y(0) + R_f s(0)[M_1 + M_2u], -R_m y(1) - R_f s(0)\}$, and $\mathbf{Q}$ is the 2×2 matrix

$$\mathbf{Q} = \begin{pmatrix} R_f(M_1 + M_2u)^2 + R_m & -R_f(M_1 + M_2u) \\ -R_f(M_1 + M_2u) & R_f + R_m \end{pmatrix}, \tag{E-64}$$

and

$$\mathbf{Q}^{-1} = \frac{1}{R_mR_f(M_1 + uM_2)^2 + R_m^2 + R_mR_f}\begin{pmatrix} R_f + R_m & R_f(M_1 + M_2u) \\ R_f(M_1 + M_2u) & R_f(M_1 + M_2u)^2 + R_m \end{pmatrix} \tag{E-65}$$

The evaluation of the expextation values of $\mathbf{Q}^{-1}$ and $\mathbf{Q}^{-1}\cdot\mathbf{W}$ using the Gaussian distribution for $P(u)$ would result in a set of equations for the

source terms, here $s(0)$, but these do not appear to be tractable analytically, so we suggest thinking about an expansion in M_2 which is proportional to the gradient of the concentration field, and this may be small. This expansion, about the mean u_0 will allow as estimate of the uncertainty in the source values inheritied from uncertainty in the observations, the model, and the wind field transporting the tracer.

Relation to Bocquet, 2005

In Bocquet's paper [1] he solves for the source $s_a(n)$ concentration relation using a variational principle quite separate from the expression given by Equation (E-56). He assumes a minimization of the entropy, thus emphasizing one's ignorance at the outset of the estimation, and seeking to improve one's knowledge through the assimilation of data. This method does not require any additional assumptions beyond saying something about the errors in the measurements–Gaussian here, but not necessarily for Bocquet, and the errors in the model. Bocquet assumes the model has no errors and that the concentration $c(\mathbf{r}, t)$ can be evaluated from the transport equations in terms of the sources $s(\mathbf{r}, t)$ via

$$c(\mathbf{r}, t) = \int d^3r' G(\mathbf{r}, t; \mathbf{r}', 0) s(\mathbf{r}', 0). \tag{E-66}$$

He then compares the measurements $y_l(t)$ with a measurement function applied to $c(\mathbf{r}, t)$ calculated in terms of the sources using this Green function. Instead, however, of minimizing the least squares quantity

$$\int_0^T dt (y_l(t) - Hc(\mathbf{r}, t))^2, \tag{E-67}$$

he selects another principle, whose Physics is not described, which relates a distribution in $s(\mathbf{r}, t)$ to an *a priori* distribution of sources via a minimum entropy principle.

E.7 Summary and Status

This appendix has summarized a number of approaches to the use of observed data in estimating the concentration of trace gases, especially CO_2, in the atmosphere/ocean system. The transport equation at the opening of this appendix is somewhat deceptive as estimates easily show: the molecular viscosity, diffusion transport coefficient, is so small, it is irrelevant. The real transport, in addition to advection by the wind, is from fluctuations in the wind $u(x,t)$. Estimating this well is at the heart of the transport of a passive tracer such as the gases we address in this report. (When the gases are reactive, everything is changed.) The general methods for data assimilation as a statistical physics problem are under development in the study of problems including the CO_2 problem addressed in this report. Over several decades the focus in numerical models of complex phenomena has been on increasing spatial and temporal resolution and adding representations of detailed physical processes. The importance of using data from observations of parts of these complex systems has been of some importance for a shorter period, and computational power may not yet be enough to do the tasks required to perform the tasks of model completion: estimate states and parameters from sparse data or testing: identify models inconsistent with the data. We do not expand on this effort here. Instead we make some observations within the challenges of estimating fluxes of transported tracer gases. The methods surveyed here fall into two broad categories:

1. The wind is given by another calculation. Typically the wind field, resolved to some spatial level, is presented as an externally derived data set. If it does not have both the mean wind and the fluctuations about the mean, it is not sufficient to provide accurate tracer transport for the reason just noted. If it does, then the ability to identify gas concentrations is likely to be washed out by the diffusion induced by wind fluctuations, but fluxes of CO_2 or other gases are more likely to be accurate. In any case, coupling one model calculation for the wind

field with another for the transport of CO_2 may be acceptable, but has the flaw of putting together two different estimation procedures with no sense of correlation between the two sets of observations or dynamical variables. Ray Weiss (SIO) reported to JASON about estimates of very, very dilute concentrations of tracer gases where the fluctuation against a zero background were reliably estimated. In this situation, perhaps putting together two different calculations might be valuable.

2. A joint model of the atmosphere/ocean coupled to the transport of CO_2 is solved along with the incorporation of information from observations of all dynamical variables: wind, surface pressure, temperature, as well as CO_2 concentrations, directly or through measurement functions (called $h(x)$ in the text). This coupling of the weather model plus the transport model requires consistency of the two, and provides a firmer basis for the estimation of CO_2 fluxes. In this regard the work of Fung and Kalnay and associates [15, 16] is very promising, and, while harder computationally, appears to be the right direction.

As discussed in Section 5 we learned about several controlled releases in the EU where using the first method there were striking successes and failures. The data assimilation method of Bouquet was used in these calculations.

We did not hear in person about the work at the European Centre for Medium-Range Weather Forecasts (ECMWF) in their GEMS program, but a review talk by Richard Egelen was available to us (Monitoring of Atmospheric CO_2 and CH_4, 2007). It uses the 4DVar variational principle described in the text (saddle path method). The work of Fung, Kalnay, Kang, Liu, and collaborators (c.f. [15, 16]) is more ambitious as it uses an ensemble data assimilation scheme for the coupled transport and weather model problem. In the latter case estimates of the mean (4DVar estimates the mode) and RMS errors about that mean are the goal.

The estimation of the horizontal wind and vertical transport processes, especially, and that along with the other dynamical variables of weather is the

best way to provide a consistent picture of CO_2 or other fluxes. We encourage NNSA and DOE to support the ensemble methods for data assimilation. They will prove useful in the CO_2 problem as well as in many other problems of interest to DOE: transport in nuclear power plants, dynamics of genetic networks, stability of power grids, where noisy data and model errors certainly play roles in estimating conditional probabilities of interesting events.

Appendix E – References

[1] Bocquet, M., "Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: Theory", *Quarterly Journal of the Royal Meteorological Society*, **131**, 2209-2223, (2005).

[2] A. C. Lorenc, *Q. J. R. Meteorol. Soc.* **112**, 1177-94 (1986).

[3] D. T. Pham, *Mon. Weath. Rev.* **129**, 1194-1207 (2001).

[4] Cohn, S. E., "An Introduction to Estimation Theory," Data Assimilation Office Note 97-01, Goddard Space Flight Center, 28 May 1997. http://gmao.gsfc.nasa.gov/pubs/docs/Cohn192.pdf

[5] H. Nijmeijer and I. M. Y. Mareels, " An Observer Looks at Synchronization," *IEEE Trans. Circuits Syst. I*, **44**, 882-890 (1997).

[6] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, 376pp, (1970).

[7] Abarbanel, H. D. I., "Effective actions for statistical data assimilation," *Physics Letters A* **373**, 4044-4048 (2009). doi:10.1016/j.physleta.2009.08.072

[8] Quinn, J. and H. D. I. Abarbanel, "State and Parameter Estimation using Monte Carlo Evaluation of Path Integrals," submitted to *Quarterly Journal of the Royal Meteorological Society*, December, 2009.

[9] Abarbanel, H. D. I., M. Kostuk, and W. Whartenby, "Data Assimilation with Regularized Nonlinear Instabilities," *Quarterly Journal of the Royal Meteorological Society* **136**, 769783 (2010).

[10] Gourdi, S. M., *et al*, "Regional-scale geostatistical inverse modeling of North American CO_2 fluxes: a synthetic data study," *Atmospheric Chemistry and Physics Discussions* **9**, 22407-22458 (2009).

[11] Judd, K. and T. Stemler, "Failures of sequential Baysian filters and the successes of shadowing filters in tracking nonlinear deterministic and stochastic systems," *Physical Review E* **79**, 066206 (2009).

[12] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*, Springer Berlin Heidelberg New York (2007).

[13] Hunt, B., E. J. Kostelich, and I. Szunyogh, "Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble transform Kalman Filter," (2007).

[14] Najm, H. N., "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics," *Annu. Rev. Fluid Mech.* **41**, 3552 (2009).

[15] Liu, J., Fung, I., Kalnay, E.,"Propagation of uncertainties from the assimilation of meteorological observations to CO2 forecast", presented at the Fall AGU meeting (2008).

[16] Kang, J., "Carbon cycle data assimilation using a coupled atmosphere-vegetation and the Local Ensemble Transform Kalman Filter", dissertation, Univ. of Maryland (2009).

F APPENDIX: Use of Tracers for Empirically Testing Models

Uncertainty in the regional estimates for GHG emissions include measurement errors of green house gas concentrations as well as the uncertainties due to the weather and transport models that are used to convert measured concentrations into emission estimates. In some cases there is an additional uncertainty due to "prior" assumptions about the distribution of sources of emission in space and time. In many circumstances the measurement errors, model uncertainties, and prior assumptions are coupled in the analysis. Consequently an "ensemble" analysis [1] is often used in which model parameters and assumptions about emission sources are varied according to some probability distribution function (PDF) to provide an indication of the spread of predicted concentrations.

For monitoring of emissions in the context of GHG agreements and treaties, it is important to obtain a high level of confidence in the uncertainty quantification for emission estimates. In particular, since treaty monitoring requires attribution of emissions to a country or region, it is very useful to validate estimates of uncertainties in weather and transport models independently of prior assumptions about distributions of emission sources.

One approach to quantification of uncertainty in weather and transport models is to use controlled release of tracer gas to provide an emission source which is well-defined in space and time. If the measurements of concentration of the tracer gas at distant locations are sufficiently accurate, weather and transport uncertainties are empirically determined independent of prior assumptions about the source distribution. Such experiments have been carried out in the 1990s [2, 4].

F.1 Release of Atmospheric Tracer Gases

Well known atmospheric tracer gases include SF_6, perflourocarbon (PFC) compounds (c.f. [2]), and chloroflourocarbons (CFCs). We will not attempt a detailed tradeoff between the various tracer gases here, rather we discuss the feasibility and utility of using one particular approach, perflourocarbon tracers (PFTs).

PFTs are inert, non-toxic, highly stable compounds consisting of 4-6 atom carbon rings [2]. They do not deplete stratospheric ozone. Examples of perflourocarbon compounds, typical background concentrations, and total atmospheric burden are given in Table 6.

Table 6: Characteristics of PFTs and Release Experiments

	PDCB	PMCH	oPDCH
Name (perflouro-)	dimethylcyclobutane	methylcyclohexane	1,2-dimethylcyclohexane
Chemistry	C_6F_{12}	C_7F_{14}	C_8F_{16}
Background level (ppqv)	3 ± 1	8 ± 1	1 ± 1
Atmospheric burden	140 metric ton	500 metric ton	71 metric ton
Release experiment (name)	BRAVO	ANATEX+ETEX+others	ANATEX+others
Amount released (kg)	1600	800 (ANATEX)	3300

Note that the amounts released in a single release experiment can sometimes be a few percent of the total atmospheric burden. Because the lifetime of these gases in the atmosphere is long (> 3000 yr), prudence must be exercised in the release of these gases. However, although the per molecule potency is large ($\sim$ 1 W m^{-2} ppb^{-1}), their low concentration in the atmosphere implies that the radiative forcing will be five orders of magnitude less than that of CO_2 (1.7 W m^{-2}). Consequently, the main concern for individual experiments is not GHG radiative forcing, but rather increase in the ambient global background levels. In this regard, we note recent reports that China has become a major emitter of PFCs [5] with emissions of order hundreds of tons per year for some PFC species. However, emissions have not been reported for some of the specific PFT gases mentioned above as possible tracer gases.

What level of downstream enhancement might be achieved in a controlled

release experiment? Consider instantaneous release of 100 kg of a PFT. The concentration in the plume at a point downwind is, in normalized units:

$$C_{PFT} = \; 3 \times 10^{-14} \times \left(\frac{M_{release}}{100 \text{ kg}}\right) \left(\frac{MW_{PFT}}{400 \text{ g/Mole}}\right)^{-1} \left(\frac{Area}{10^5 \text{ km}^2}\right)^{-1} \times \left(\frac{h_{mix}}{1.5 \text{ km}}\right)^{-1} \left(\frac{Mf_{air}}{45 \text{ Mole/m}^3}\right)^{-1}$$

i.e. about 10 times the ambient background concentration and over 30 times the typical error for concentration measurements of PFTs (10^{-15} or 1 ppqv). Careful measurements may do an order of magnitude better, implying percent-level accuracy in measurement of concentrations in the downwind plume. These estimates are meant to be order-of-magnitude estimates of feasibility, not a detailed calculation of an actual release experiment.

F.2 Continental-Scale Release Experiments in the Past

Several experiments of controlled release of tracer gases have been carried out in the past on continental scales. These include the *Across North America Tracer Experiment* (ANATEX) [6] and the *European Tracer Experiment* (ETEX)[3]. See Table 7.

Table 7: Release Experiments

Name	Year	Location	Number	PFT Species	Total (tonne)	Sampling
ANATEX	1987	North America	30	PMCH,oPDCH PTCH	7.4	77 surface, 5 tower aircraft
ETEX	1994	Europe	2	PMCH,PMCP	0.83	168 surface

The ANATEX experiment was carried out in 1987 and consisted of 30 individual releases of PFTs from two sites. Samples were taken at approximately 77 surface-level sites and 5 towers, as well as aircraft measurements. Figure 48 shows a simulation of the release plume for one of the ANATEX experiments.

The ETEX experiment was performed in 1994 and consisted of two releases from western France of two PFTs. The release was sampled by 168 stations.

Figure 49 shows a simulation of the release plume for the ETEX experiment. Results of analysis of the ETEX experiments are discussed in the main body of this report and also in [7]. These results indicate the importance of controlled release experiments for quantitative assessment of modeling errors.

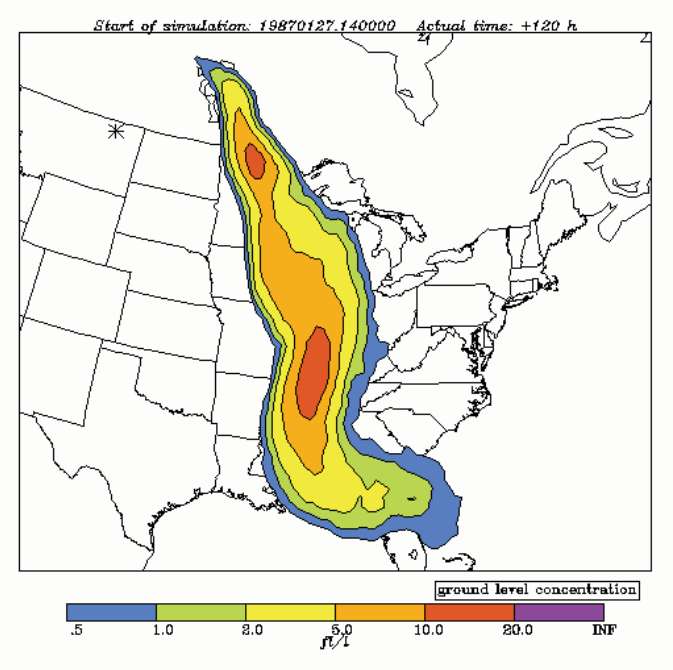


Figure 48: Simulation of the plume 5 days after an ANNATEX release. Asterisk marks the release point. (http://zardoz.nilu.no/ andreas/flexpart.html)

F.3 Recommended capabilities for tracer gas release

Given the importance of atmospheric transport in deriving GHG emissions from GHG concentration measurements, it is surprising that major experiments with controlled release of tracers have not occurred since the 1990s. We recommend that DOE develop the capability to carry out "challenge" campaigns in which a controlled release of a tracer gas is followed by extensive surface and air sampling to measure tracer gas concentrations. Weather-transport and inversion codes would then be challenged to analyze the measurements. Much of the required capability already exists within the DOE to execute such a challenge campaign. Additional capability exists within

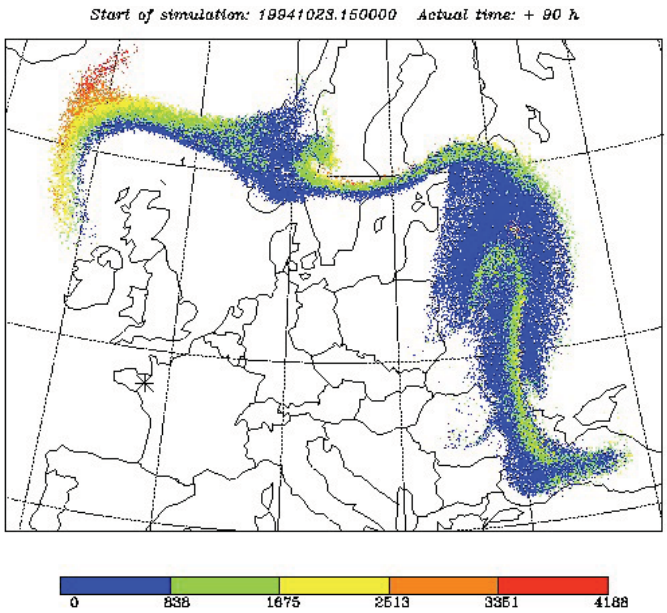


Figure 49: Simulation of the plume 89 hours after an ETEX release. Asterisk marks the release point. Colors indicate plume particle height (meter). (http://zardoz.nilu.no/ andreas/flexpart.html)

NOAO and significant expertise also resides in Europe.

Appendix F – References

[1] Kang, J., "Carbon cycle data assimilation using a coupled atmosphere-vegetation and the Local Ensemble Transform Kalman Filter", dissertation, Univ. of Maryland (2009).

[2] Watson et al., "The Atmospheric Background of Perflourocarbon Compounds Used as Tracers", *Env. Sci. and Tech.*, **41**, 6909 (2007)

[3] Nodop, K. et al., "The Field Campaigns of the European Trace Experiment (ETEX): Overview and Results", *Atmospheric Environment*, **32**, 4905-4108 (1998).

[4] Stohl, A., "Computation, Accuracy and Applications of Trajectories – A Review and Bibliography", *Atmospheric Environment*, **32**, 947-966 (1998).

[5] Saito, T. et al., "Large Emissions of Perfluorocarbons in East Asia Deduced from Continuous Atmospheric Measurements", Environ. Sci. Technol., **44**, 4089-4095 (2010).

[6] R.R. Draxler, R. Dietz, R.J. Lagomarsino and G. Start, "Across North America tracer experiment (ANATEX): Sampling and analysis", Atmospheric Environment. Part A., Volume 25, Issue 12, Pages 2815-2836 (1991).

[7] M. Krysta1, M. Bocquet, and J. Brandt, "Probing ETEX-II data set with inverse modeling", *Atmos. Chem. Phys.* **8**, 3963-71 (2008).

G APPENDIX: Selected Acronyms

ABET	Accreditation Board for Engineering and Technology
AIRCOA	Autonomous Inexpensive Robust CO2 Analyzer
AIRS	Atmospheric Infrared Sounder
AMS	Accelerator Mass Spectrometer
AMU	Atomic Mass Unit
ANATEX	Across North America Tracer Experiment
ARGO	Array for Real-time Geostrophic Oceanography
ASCENDS	Active Sensing of CO2 Emissions over Nights, Days, and Seasons
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
CCS	Carbon Capture and Storage
CONUS	Continental United States
CTBT	Comprehensive nuclear-Test-Ban Treaty
CFC	Chlorofluorocarbon
CIA	Central Intelligence Agency
CRDS	Cavity Ring-Down Spectroscopy
DOE	Department of Energy
ECMWF	European Centre for Medium-Range Weather Forecasts
EDGAR	Emissions Database for Global Atmospheric Research
EIA	Energy Information Administration
ETEX	European Tracer Experiment
EU	European Union
FFDAS	Fossil Fuel Data Assimilation System
FTS	Fourier Transform Spectroscopy
GCM	Global Climate Model
GEO	Geosynchronous Orbit
GHG	Greenhouse Gas
GOSAT	Greenhouse gas Observing SATellite
GPS	Global Positioning System
GSD	Ground Sampling Distance
HCF	Hydrofluorocarbon
HCFC	Hydrochlorofluorocarbon
HVAC	Heating, Ventilating, and Air Conditioning
IGCC	Integrated Gasification Combined Cycle
IPCC	Intergovernmental Panel on Climate Change
IR	Infrared
LED	Light Emitting Diode
LEO	Low Earth Orbit
LIDAR	Light Detection and Ranging

LLNL	Lawrence Livermore National Laboratory
MEA	Methyl-Ethanol-Amine
MODIS	Moderate Resolution Imaging Spectroradiometer
NAMA	Nationally Appropriate Mitigation Action
NCAR	National Center for Atmospheric Research
NDIR	Non-Dispersive Infrared
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
NTM	National Technical Means
OCO	Orbiting Carbon Observatory
PFC	Perfluorocarbon
PFT	Perfluorocarbon tracer
ppb	parts per billion
ppm	parts per million
RACCOON	Regional Atmospheric Continuous CO2 Network
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CartograpHY
STP	Standard Temperature and Pressure
SWIR	Short-wave Infrared
TIR	Thermal Infrared
TOPEX	Topography Experiment
UNFCCC	United Nations Framework Convention on Climate Change
V-PDB	Vienna Pee-Dee Belemnite
V-SMOW	Vienna Standard Mean Ocean Water
VNIR	Visible Near Infrared
WMO	World Meteorological Organization